

# **The effects of Tma20 and Tma22 on gene expression and telomere function**

(1 volume)

**Victoria Torrance**

Thesis submitted for the degree of Doctor of Philosophy

Institute for Cell and Molecular Biosciences

Newcastle University

September 2017



## Abstract

Translation Machinery Association 20 (Tma20) and Tma22 function as interacting partners and are highly conserved. *TMA20* and *TMA22* are among a diverse range of yeast genes, which have been shown to decrease fitness of cells with *cdc13-1* induced telomere defects. Deletions of nonsense mediated decay genes (NMD) also increase fitness of *cdc13-1* and it is thought that this is partially due to the higher levels of Stn1 in *NMD* null strains.

Genetic interaction studies show that *TMA20* and *TMA22* function in the same pathway as NMD genes, but in parallel pathways to DNA damage genes, to affect fitness of *cdc13-1*. This led us to hypothesise that *TMA20* and *TMA22* also affect fitness of *cdc13-1* by increasing levels of Stn1. Consistent with this, we found that *tma20* $\Delta$  strains do indeed have increased levels of Stn1. In *Drosophila* Tma20 and Tma22 regulate the expression of genes that contain upstream ORFs (uORFs) by promoting translation re-initiation. However we found no evidence that Tma20 and Tma22 promote translation re-initiation in yeast, as they do in *Drosophila*.

Interestingly though, we observed that *STN1* has a uORF and also an ORF that overlaps with the main coding sequence of *STN1* which we refer to as an oORF. We demonstrate that the increase in expression of *STN1* that occurs upon deletion of *TMA20* is dependant on the ORF that overlaps with the main coding sequence of *STN1*, rather than the uORF. We show that the oORF of *STN1* serves as an important regulatory element, which dramatically reduces levels of Stn1. Fitness of *cdc13-1* is substantially increased in strains that contain a point mutation in the initiation codon of the oORF. The human homolog of *STN1* also has an oORF, suggesting this mechanism of regulating levels of crucial telomere proteins may be conserved.

## Acknowledgments

I would like to thank my supervisor, David Lydall for accepting me as a PhD student, for his support and invaluable guidance and for giving me the opportunity to present my work at two conferences. I would like to thank all the members of the Lydall laboratory (Marta, Joana, Marion, Alan, Conor, Ben, Cameron, Eva, Sid, Greg and Kate) who have been there during my time in the lab, for all their help and friendship throughout my studies. I am especially grateful to Marta, Joana, Marion and Greg who have always been helpful when troubleshooting as well as Conor who gave me a great deal of help and encouragement with my programming. I would also like to thank to members of the Perkins laboratory especially Iglika, H  l  ne and Ling-I who continued to provide me with advice whenever I have needed it since I finished my master's project with them. I extend my gratitude also to Neil Perkins as my second supervisor for allowing me to complete my MRes project in his laboratory and for being such a helpful supervisor during this time. I am grateful to my internal assessors Jeremy Brown and Nick Watkins for providing constructive criticism and advice at each annual assessment. Finally I'm very grateful for the BBSRC for funding my studies.





## Table of Contents

List of Figures.....	xi
Yeast Nomenclature.....	xiv
List of Abbreviations.....	xv
 <b>Chapter 1. Introduction .....</b>	<b>1</b>
<b>1.1 Telomeres protect the linear ends of chromosomes .....</b>	<b>1</b>
1.1.1 Organisation of telomere DNA sequences .....	1
1.1.2 Telomere replication and elongation.....	1
<b>1.2 Protein component of the telomere.....</b>	<b>5</b>
1.2.1 CST binds single stranded DNA at telomeres .....	5
1.2.2 Telomere proteins associated with double stranded DNA.....	6
<b>1.3 Telomere uncapping leads to cell cycle arrest .....</b>	<b>8</b>
<b>1.4 High throughput screens uncover genes involved in telomere biology</b>	<b>11</b>
<b>1.5 Tma20, Tma22 and Tam64 .....</b>	<b>11</b>
<b>1.6 MCT-1 and DENR function in translation.....</b>	<b>15</b>
<b>1.7 Translation is the synthesis of proteins from mRNAs .....</b>	<b>15</b>
<b>1.8 Translation Initiation.....</b>	<b>19</b>
<b>1.9 Some transcripts contain uORFs .....</b>	<b>23</b>
1.9.1 uORF mediated translational control .....	23
1.9.2 uORFs facilitate expression of some genes in response to stress .....	27
1.9.3 Specific examples of translation regulation by uORFs .....	27
1.9.4 Tma20 and Tma22 promote translation of genes that contain uORFs....	32
<b>1.10 Translation termination and ribosome recycling .....</b>	<b>32</b>
<b>1.11 Translation in conditions of stress .....</b>	<b>36</b>
<b>1.12 Nonsense mediated Decay .....</b>	<b>36</b>
<b>1.13 Aims .....</b>	<b>40</b>
 <b>Chapter 2. Materials and Methods .....</b>	<b>41</b>
<b>2.1 Recipes for yeast and bacterial media.....</b>	<b>41</b>

2.1.1	Yeast .....	41
2.1.2	E-coli .....	41
<b>2.2</b>	<b>Yeast genetic methods .....</b>	<b>41</b>
2.2.1	Mating .....	41
2.2.2	Sporulation .....	41
2.2.3	Tetrad analysis .....	42
2.2.4	Gene deletion and epitope tagging .....	42
2.2.5	High efficiency Lithium acetate transformation .....	42
2.2.6	Transformation of plasmids into yeast (One-step transformation) .....	42
2.2.7	Isolation of plasmids from yeast .....	43
2.2.8	Growth assays .....	43
2.2.9	Microscopy .....	43
<b>2.3</b>	<b>Molecular biology methods .....</b>	<b>43</b>
2.3.1	Extraction of Genomic DNA from yeast (for PCR) .....	43
2.3.2	Extraction of Genomic DNA from yeast (for Southern Blot) .....	43
2.3.3	Plasmid digestions .....	44
2.3.4	Polymerase chain reaction (PCR) .....	44
2.3.5	Gel electrophoresis .....	45
2.3.6	<i>In vivo</i> cloning .....	45
2.3.7	One step Isothermal DNA assembly (Gibson Assembly) .....	45
2.3.8	Plasmid transformation into <i>E-coli</i> .....	46
2.3.9	Plasmid isolation from <i>E-coli</i> .....	46
2.3.10	Luciferase assays .....	46
2.3.11	TCA protein extraction .....	46
2.3.12	Western Blotting .....	47
2.3.13	RNA isolation .....	47
2.3.14	qPCR .....	48
2.3.15	Southern Blotting .....	49
<b>2.4</b>	<b>Bioinformatics .....</b>	<b>50</b>
2.4.1	Obtaining TL sequences .....	50
2.4.2	Scoring initiation contexts .....	50
2.4.3	Analysis of TLs .....	50
<b>Chapter 3. TMA20 and TMA22 act in the same pathway as NMD genes to increase fitness of <i>cdc13-1</i> cells .....</b>		<b>51</b>

<b>3.1 Tma20 and Tma22 act as interacting partners to decrease fitness of <i>cdc13-1</i> and <i>stn1-13</i>.....</b>	<b>51</b>
<b>3.2 Tma20 does not affect telomere length .....</b>	<b>56</b>
<b>3.3 <i>HCR1</i> strongly reduces fitness of <i>cdc13-1</i> cells and acts in a different pathway as <i>TMA20</i> and <i>TMA22</i>.....</b>	<b>59</b>
<b>3.4 <i>cdc13-1 hcr1Δ</i>, but not <i>cdc13-1 tma20Δ</i>, is suppressed by overexpression of <i>RLI1</i> .....</b>	<b>63</b>
<b>3.5 <i>TMA20</i> and <i>TMA22</i> act in the same pathway as <i>NMD</i> to affect fitness of <i>cdc13-1</i>.....</b>	<b>67</b>
<b>3.6 Evidence that increase in levels of stop codon read-through does not suppress <i>cdc13-1</i> .....</b>	<b>71</b>
<b>3.7 <i>TMA20</i> and <i>TMA22</i> act in parallel pathways to DNA damage genes .....</b>	<b>75</b>
<b>3.8 Mutations in conserved phosphorylation sites affect function of the Tma20 .....</b>	<b>79</b>
<b>3.9 No evidence that Tma20 is up regulated or phosphorylated following telomere uncapping .....</b>	<b>83</b>
<b>3.10 Discussion .....</b>	<b>86</b>
3.10.1 Interacting partners, Tma20 and Tma22 have a novel role in telomere biology	86
3.10.2 Decreased recycling of post termination ribosomes improves <i>cdc13-1</i> fitness	86
3.10.3 Tma20 and Tma22 function in the same pathway as NMD to decrease fitness of <i>cdc13-1</i> .....	87
3.10.4 Increased stop codon read-through reported in <i>tma20Δ</i> , <i>nmd2Δ</i> and <i>hcr1Δ</i> does not explain the increase in <i>cdc13-1</i> fitness observed in these strains	88
3.10.5 No evidence to suggest that Tma20 actively responds to <i>cdc13-1</i> induced DNA damage .....	89

<b>Chapter 4. No Evidence that Tma20 and Tma22 promote translation re-initiation as in <i>Drosophila</i> and mammalian cells.....</b>	<b>90</b>
<b>4.1 Identification of genes with uORFs .....</b>	<b>90</b>
<b>4.2 Genes with uORFs have lower translation rates and uORFs are underrepresented in the genome. ....</b>	<b>94</b>
<b>4.3 uORFs have weaker initiation contexts than main ORFs .....</b>	<b>98</b>

4.4	uORFs are longer and terminate further from main ORFs than expected.	101
4.5	New luciferase reporters created to measure translation <i>in vivo</i> .....	104
4.6	uORFs, as expected, inhibit the expression of main ORFs to varying extents. ....	108
4.7	No evidence that Tma20 or Tma22 promote translation re-initiation in yeast.....	117
4.8	Evidence suggests that Tma20 and Tma22 decreases transcript stability, as demonstrated using alleles of <i>BRE4</i> .....	121
4.9	All three uORFs of <i>ALD3</i> are required to promote translation.....	129
4.10	Discussion.....	133
4.10.1	Regulation of gene expression by uORFs.....	133
4.10.2	No evidence that Tma20 and Tma22 promote translation re-initiation	134
4.10.3	Interaction of Tma20 and Tma22 with the ribosome has multiple consequences for translation .....	134
Chapter 5.	Expression of Stn1 is tightly controlled by an oORF.....	139
5.1	<i>STN1</i> upstream open reading frames decrease reporter gene expression	139
5.2	<i>STN1</i> mutations strongly increase <i>STN1</i> expression from the native locus. ....	142
5.3	<i>STN1-101</i> and <i>STN1-102</i> improve fitness of cells with <i>cdc13-1</i> telomere defects. ....	146
5.4	<i>TMA20</i> and <i>NMD2</i> reduce Stn1 levels.....	149
5.5	<i>TMA20</i> and <i>NMD2</i> reduce Stn1 levels via the overlapping uORF. ....	152
5.6	Genetic evidence supports the idea that <i>TMA20</i> decreases fitness of <i>cdc13-1</i> by decreasing levels of Stn1 .....	155
5.7	<i>STN1-102</i> and <i>STN1-101</i> do not affect sensitivity to HU.....	159
5.8	Discussion .....	162
5.8.1	Stn1 levels are tightly controlled by an oORF .....	162
5.8.2	Degradation of Stn1 by NMD .....	163
5.8.3	Tma20 and Tma22 decrease fitness of <i>cdc13-1</i> by decreasing Stn1 levels	164
5.8.4	Tma20 and Tma22 regulate expression of genes with oORFs.....	164
Chapter 6.	General Discussion .....	166

6.1 Tma20 and Tma22 decrease fitness of <i>cdc13-1</i> by reducing levels of Stn1	168
6.2 Genetic interactions between <i>TMA20</i> and <i>NMD2</i> .....	166
6.3 Levels of Stn1 are controlled by an oORF.....	167
6.4 Multiple ways that Tma20 and Tma20 can affect expression of genes that contain uORFs or oORFs.....	168
Appendix A. Strains.....	170
Appendix B. Plasmids .....	176
Appendix C. Primer sequences .....	187
References .....	193

## List of Figures

Figure 1 - Telomere replication and elongation.....	3
Figure 2 - Structure of a yeast telomere.....	7
Figure 3- Telomere uncapping leads to cell cycle arrest.....	9
Figure 4- Domain organisation of Tma20, Tma22 and Tma64 in yeast and humans	13
Figure 5 – Translation initiation, elongation, termination and ribosome recycling .....	17
Figure 6 – Translation initiation .....	21
Figure 7 – uORF mediated translational control.....	25
Figure 8 - Translation re-initiation controls the expression of <i>GCN4</i> .....	30
Figure 9 - Translation termination and ribosome recycling .....	34
Figure 10 – Nonsense mediated mRNA decay .....	38
Figure 11 - Tma20 and Tma22 act as interacting partners to decrease fitness of <i>cdc13-1</i> and <i>stn1-13</i> .....	54
Figure 12 - Tma20 does not affect telomere length .....	58
Figure 13 – <i>HCR1</i> strongly improves fitness of <i>cdc13-1</i> cells and acts in a different pathway as <i>TMA20</i> and <i>TMA22</i> .....	61
Figure 14 - <i>cdc13-1 hcr1Δ</i> , but not <i>cdc13-1 tma20Δ</i> , is supressed by overexpression of <i>RLI1</i> .....	65
Figure 15 - <i>TMA20</i> and <i>TMA22</i> act in the same pathway an <i>NMD</i> to affect fitness of <i>cdc13-1</i> .....	69
Figure 16 - Evidence that increase in levels of stop codon read-through does not supress <i>cdc13-1</i> .....	73
Figure 17 - <i>TMA20</i> and <i>TMA22</i> act in parallel pathways to DNA damage genes .....	77
Figure 18 - Mutations in conserved phosphorylation sites affect function of the Tma20 .....	81

Figure 19 - No evidence that Tma20 is upregulated or phosphorylated following telomere uncapping .....	85
Figure 20 – Identification of genes with uORFs .....	92
Figure 21 - Genes with uORFs have lower translation rates and uORFs are underrepresented in the genome. ....	96
Figure 22 - uORFs have weaker initiation contexts than the main ORFs.....	99
Figure 23 - uORFs are longer and further from main ORFs than expected .....	102
Figure 24 - New luciferase reporters created to measure translation <i>in vivo</i> .....	106
Figure 25 - 22 genes with uORFs selected to explore the role of Tma20 and Tma22 in translation .....	111
Figure 26 - uORFs, as expected, inhibit the expression of main ORFs to varying extents. ....	114
Figure 27- uORFs that reside close to the main ORF correlate with reduced translation .....	115
Figure 28 - No evidence that Tma20 or Tma22 promote translation re-initiation in yeast .....	119
Figure 29- <i>BRE4</i> expression is decreased by a uORF .....	123
Figure 30- Tma20 and Tma22 decrease expression of <i>BRE4</i> by decreasing transcript levels .....	127
Figure 31 - All three uORFs of <i>ALD3</i> are required to promote translation.....	131
Figure 32 - Proposed interactions between homologs of Tma20 and Tma22 with the post-termination ribosome .....	137
Figure 33 - <i>STN1</i> upstream open reading frames decrease reporter gene expression .....	141
Figure 34 - <i>STN1</i> mutations strongly increase <i>STN1</i> expression from the native locus .....	144



Figure 35- <i>STN1-101</i> and <i>STN1-102</i> improve fitness of cells with <i>cdc13-1</i> telomere defects.....	147
Figure 36 - <i>TMA20</i> and <i>NMD2</i> reduce Stn1 levels.....	151
Figure 37 - <i>TMA20</i> and <i>NMD2</i> reduce Stn1 levels via the overlapping uORF.....	153
Figure 38- Genetic evidence supports the idea that <i>TMA20</i> decreases fitness of <i>cdc13-1</i> by decreasing levels of Stn1 .....	157
Figure 39- <i>STN1-102</i> and <i>STN1-101</i> do not affect sensitivity to HU.....	160
Figure 40 – Comparison of yeast and human Stn1 sequences .....	164

## Yeast Nomenclature

<i>YFG</i>	Wild-type gene
<i>yfgΔ</i>	Gene has been deleted
	Denotes that the gene <i>yfg</i> has been disrupted by the integration of
<i>yfg::XXX</i>	gene XXX
Yfg	Denotes that this is a protein
<i>yfg-m</i>	Denotes an alleles of a WT <i>YFG</i> gene (e.g. <i>cdc13-1</i> )

## List of Abbreviations

aa	Amino acid
ALT	Alternative lengthening of telomeres
bp	base pairs (nucleotides)
BSE	Bovine spongiform encephalopathy
CDS	coding sequence of a gene
CRAC	UV crosslinking and analysis of cDNA
CST	Cdc13/Stn1/Ten1
DDR	DNA damage response
DEPC	Diethyl pyrocarbonate
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
DSB	Double-strand break
dsDNA	Double-stranded DNA
DTT	DT-dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
eIF	Eukaryotic initiation Factor
ESM	Enriched sporulation medium
g	Gramme
GDP	Guanosine diphosphate
GEF	Guanosine exchange factor
GFP	Green fluorescent protein
GO	Gene ontology
GTP	Guanosine-5'-triphosphate (GTP)
HR	Homologous Recombination
HRP	Horseradish peroxidase
HU	Hydroxyurea
IQR	Interquartile range
IRES	Internal ribosome entry site

ISR	Integrated stress response
kb	Kilobases (nucleotides)
kDa	KiloDaltons
L	Litre
LB	Lysogeny broth
M	Molar
MAPK	Mitogen activated protein kinase
Met-tRNA <sub>i</sub> <sup>Met</sup>	Methionine charged initiator tRNA
mg	Milligrammes
mL	Millilitre
MMS	Methyl methanesulfonate
mRNA	Messenger RNA
MRX	Mre11/Rad50/Xrs2 complex
NAD	Nicotinamide adenine dinucleotide
ng	Nanogrammes
NHEJ	Non-homologous end joining
NMD	Nonsense mediated mRNA decay
oORF	Overlapping open reading frame
ORF	Open reading frame
PBS	Phosphate-buffered saline pH 7.5
PBST	PBS with 0.1% (v/v) Tween-20
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PIC	Pre Initiation complex
PP	Pol $\alpha$ /primase
PRF	Programmed ribosomal frameshift
PTC	Premature termination codon
PUA	Pseudouridine synthase and archaeosine transglycosylase domain
QFA	Quantitative fitness analysis

qPCR	Quantitative polymerase chain reaction
qRT-PCR	Quantitative reverse-transcriptase polymerase chain reaction
RNA	Ribonucleic acid
RPA	Rfa1/Rfa2/Rfa3
RT	Reverse transcriptase
SD media	Synthetic Drop-Out Media
SDS	Sodium dodecyl sulphate
SEM	Standard error of the mean
SGD	Saccharomyces Genome Database
ssDNA	Single-stranded DNA
TBE	Tris-borate EDTA pH 8.3 (Bio-Rad)
TC	Ternary complex
TCA	Trichloroacetic acid
TE buffer	Tris-EDTA buffer
TL	Transcript leader
tRNA	Transfer RNA
uORF	Upstream open reading frame
URS	Upstream regulatory sequence
<i>YFG</i>	Your favourite gene (any gene)
YPD	Yeast extract, peptone, dextrose
µg	Microgrammes
µL	Microlitre



# Chapter 1. Introduction

## 1.1 Telomeres protect the linear ends of chromosomes

Telomeres are repetitive sequences at the end of chromosomes which function to protect the genetic information contained within linear chromosomes. Telomere capping proteins are vital to the integrity of telomeres since an uncapped telomere resembles one half of a double strand break (DSB), which is a potent inducer of the DNA damage response. The DNA damage response arrests cell cycle to allow for repair and if damage is too severe to be repaired, apoptosis is induced.

### 1.1.1 Organisation of telomere DNA sequences

The termini of chromosomes in both mammals and yeast are composed of single stranded G-rich overhangs which are approximately 12–14 nucleotides in yeast (Figure 2), extending to 50-100 nucleotides in S-phase and 100-280 nucleotides in mammals (Larrivée et al., 2004, Wellinger et al., 1993, Webb et al., 2013). The double stranded component of telomeres consists of tandem repeats which are TG<sub>1-3</sub> in yeast and TTAGGG in humans (McEachern and Blackburn, 1994). Yeast telomeres also contain middle repetitive sub-telomeric elements known as X repeats and Y' elements, with X repeats residing closer to the centromere than Y' elements. X elements are heterogeneous and present in all chromosomes whereas Y' elements are found in approximately half of all chromosomes and exist in 1-4 tandem repeats (Mondoux and Zakian, 2007, Kupiec, 2014). Mammalian telomeres are considerably longer than yeast telomeres (5,000-15,000 base pairs in humans compared with 225-375 base pairs in yeast), however the majority of mammalian telomeres in somatic cells progressively shorten with each cell division whereas yeast telomere length is stably maintained (Larrivée et al., 2004).

### 1.1.2 Telomere replication and elongation

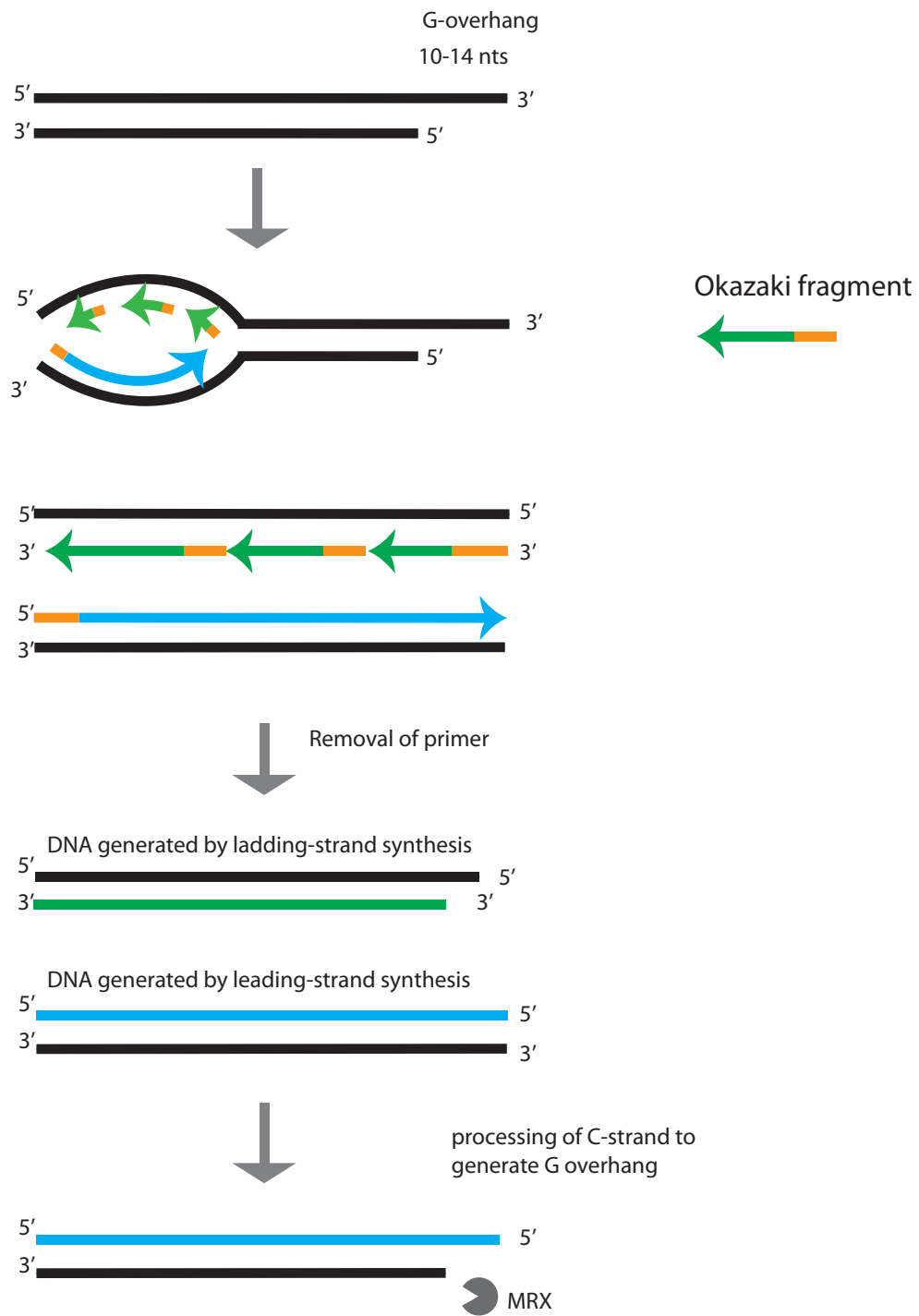
DNA replication machinery is unable to completely replicate the linear ends of DNA leading to progressive shortening of the chromosome termini that occurs with every cell division (Watson, 1972) (Olovnikov, 1973). DNA replication on the G-strand (lagging strand) occurs discontinuously by short Okazaki fragments that are ligated

together to form a continuous strand whereas synthesis of the C-strand (leading strand) occurs continuously (Figure 1a). Loss of DNA occurs on the G-rich lagging strand as a result of the so called 'end replication problem' where removal of the RNA primer on the most terminal Okazaki fragment results in a ssDNA overhang (Figure 1a). Loss of DNA is more profound on the C-rich leading strand and occurs because the blunt ends that result from leading strand replication must again be resected to provide the ssDNA required for both CST binding and telomerase mediated elongation (Figure 1a).

The progressive shortening of telomeres is counteracted by telomerase, a reverse transcriptase with an integral RNA component that adds DNA sequence repeats to the 3' end of the chromosome. Telomeres of germ and stem cells, as well as the majority of cancerous cells, are maintained by telomerase, allowing these cells to avoid senescence, which would normally occur when telomeres become critically short. Telomerase in yeast is comprised of 3 protein subunits (Est1, Est2 and Est3) and TLC1, the RNA template (Lingner et al., 1997b, Lingner et al., 1997a). Est2 is the reverse transcriptase subunit whereas Est1 and Est3 are accessory factors (Lingner et al., 1997b, Lingner et al., 1997a). In a 5' to 3' direction telomerase adds repeat sequences to the G-strand, using the integral RNA template, TLC1, as a primer (Figure 1b) (Singer and Gottschling, 1994). Primase-Pol  $\alpha$  (PP) complex then synthesises the C-strand by copying the extended G-tail produced by telomerase (Figure 1b)(Qi and Zakian, 2000) (Adams Martin et al., 2000).



A



B

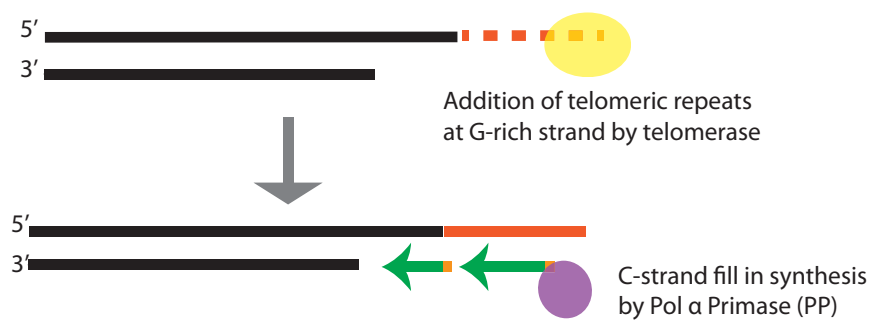


Figure 1. Telomere replication and elongation. Replication of the G-strand occurs discontinuously by Okazaki fragments and results in double stranded DNA that has a g-overhang. Replication of the C-strand occurs continuously and results in blunt ended double stranded DNA that is resected by MRX to produce a g-overhang. Telomerase adds DNA repeats to the g-strand and Pol  $\alpha$ /primase (PP) complex synthesises the C strand using the g strand as a template.

## 1.2 Protein component of the telomere

### 1.2.1 CST binds single stranded DNA at telomeres

Cdc13, Stn1, Ten1 together comprise the conserved CST complex that protects the chromosome ends, regulates telomerase activity and has structural homology to the RPA complex (Kupiec, 2014). RPA is involved in replication and repair and promiscuously binds ssDNA (Kupiec, 2014). The CST complex, on the other hand has high affinity for the G rich telomeric ssDNA (Gelinas et al., 2009). The CST complex is an essential capping component of telomeres, required to prevent the accumulation of ssDNA, activation of the checkpoint response and eventual cell cycle arrest at G2/M (Lei et al., 2004) (Grandin et al., 2001a). Although said to be part of a complex and shown to bind to one another, members of CST also have some independent functions, with Stn1 and Ten1 able to perform some capping functions in the absence of Cdc13 (Holstein et al., 2014). Cdc13 recruits telomerase through binding to its Est1 subunit, whereas Stn1 inhibits telomerase, by competitive interaction with Cdc13 (Puglisi et al., 2008, Nugent et al., 1996a). CST also facilitates C-strand fill-in synthesis by the Primase-Pol  $\alpha$  (PP) complex, as demonstrated in *Candida glabrata* and humans (Lue et al., 2014) (Ganduri and Lue, 2017). C-strand fill-in synthesis is mediated interactions between Stn1 and Pol12, the regulatory subunit of PP (Lue et al., 2014). A more general role for CST, particularly Stn1 in DNA replication has also been proposed (Chastain et al., 2016). In human cells Stn1 is thought to facilitate restart of stalled replication forks after replication stress, by promoting recruitment of Rad51 (Chastain et al., 2016). Stn1 also affects replication fork progression in yeast as cells overexpressing Stn1 fail to inhibit replication fork progression when treated with DNA damaging agents (Gasparyan et al., 2009).

The mammalian CST complex is comprised of Ctc1, Stn1 and Ten1 and also binds ssDNA at telomeres, function in chromosome end protection and telomerase recruitment (Price et al., 2010). However unlike yeast telomeres, mammalian telomeres are also protected by the Sheltin complex (Chen et al., 2012, Chen et al., 2013). Sheltin is composed of TRF1, TRF2, TIN2, RAP1, POT1 and TPP1 and binds to DNA through TRF1, TRF2 and POT1 where TRF1 and TRF2 bind double stranded DNA while POT1 binds to the ssDNA G-strand. Pot1 was originally described as an ortholog of Cdc13, however Cdc13 and Pot1 do not share homology and bind DNA by different mechanisms (Lei et al., 2004).

### *1.2.2 Telomere proteins associated with double stranded DNA*

yKu complex is a conserved DNA binding heterodimer which is comprised of subunits yKu70 and yKu80, both mediators of non-homologous end joining (NHEJ) (Kupiec, 2014). Since one of the functions of telomeres is to protect the ends of the DNA from being targeted for 'repair' by NHEJ it is surprising that yKu also binds to the double stranded component of telomeres (Figure 2). Similarly to the CST complex yKu also protects chromosome ends from exonuclease degradation, which is demonstrated by the fact that cells defective in yKu70 accumulate ssDNA which results in a cell cycle arrest (Maringele and Lydall, 2002b). The yKu complex also promotes telomerase recruitment to the telomeres via an interaction with TLC1 (Peterson et al., 2001). The double stranded TG<sub>1-3</sub> repeats are bound by Rap1, which regulates telomere length by a feedback mechanism through its association with Rif1 and Rif2 (Figure 2).

(Marcand et al., 1997). Long telomeres bind more Rap1 that in turn bind more Rif1 and Rif2, which negatively regulate telomere length. Short telomeres bind less Rap1 and thus are preferentially elongated by telomerase (Marcand et al., 1997). Rap1 and Rif2 protect the chromosome ends from degradation by nucleases and NHEJ (Marcand et al., 2008) whereas Rif1 is thought to assist the telomere protection function of the CST complex (Anbalagan et al., 2011). The SIR complex is comprised of Sir2, Sir3 and Sir4 and forms a silencing complex through interaction with histones (Kupiec, 2014).

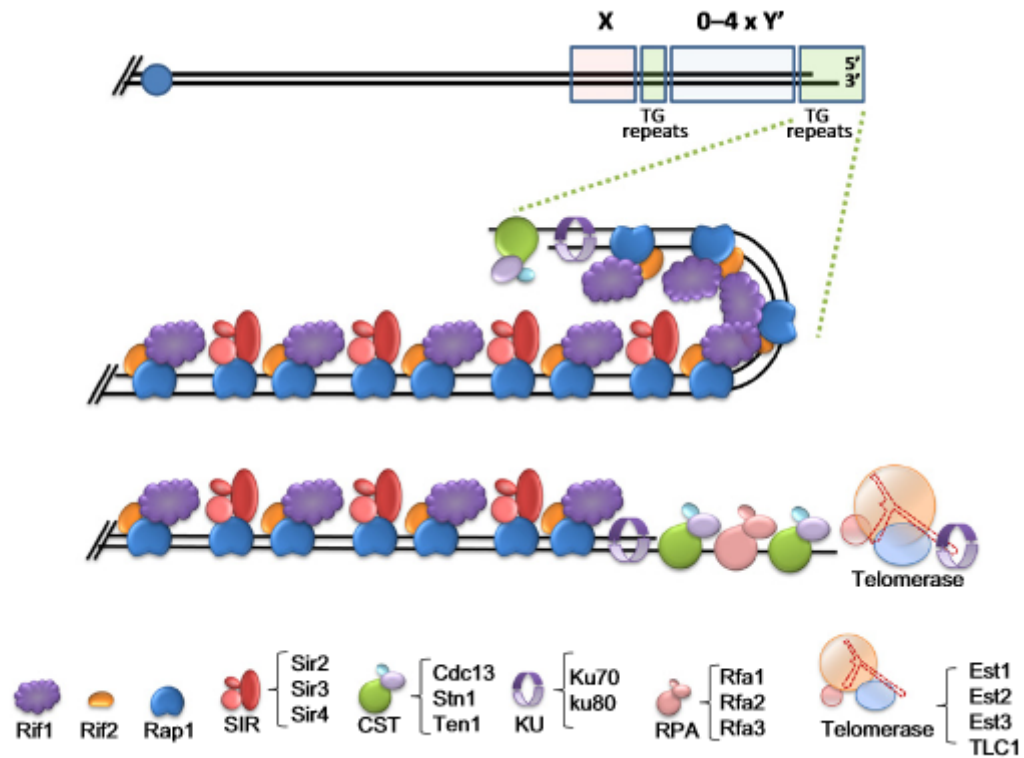


Figure 2. Structure of a yeast telomere. Schematic representation of a yeast telomere, showing the X and Y' sequences and the internal and terminal TG overhangs. Rap1 binds the telomeric repeats; and Rif1, Rif2, and the SIR proteins bind to Rap1. The Ku heterodimer binds to telomeric double strand DNA, and the CST complex binds the terminal ssDNA G-tails. Telomerase is recruited to telomeres. Figure and text taken from (Kupiec, 2014)

### **1.3 Telomere uncapping leads to cell cycle arrest**

Telomere uncapping can occur as a consequence of the disruption in protein level of any of the capping proteins, or in human cells as a consequence of telomere repeats becoming critically short (Wellinger, 2010). When telomeres become uncapped they are targeted by DSB repair machinery in a manner comparable to the targeting of actual DSBs by DSB repair machinery (Figure 3). One of the initial responses to DSBs and uncapped telomeres is resection. At DSBs this is mediated by MRX (Rad50, Mre11 and Xrs2), which surprisingly inhibit high levels of resection at telomeres (Foster et al., 2006). Resection at uncapped telomeres is carried out by Exo1 and Dna2 nucleases, the latter of which functions in combination with the helicase Sgs1 (Ngo et al., 2014). YKu complex binds to DSBs, promoting NHEJ as an alternative repair pathway to HR by inhibiting resection. Resection creates ssDNA, a target of RPA which recruits Rad51 recombinase and Rad52, both mediators of HR repair (Wellinger, 2010). RPA is also important in the promotion of cell cycle arrest since it activates checkpoint kinase Mec1, Mec1 binding partner Ddc2 and the 9-1-1 complex (Ellison and Stillman, 2003). The 9-1-1 complex, comprised of Mec3, Ddc1 and Rad17 is loaded to sites of damage by the so called clamp loader Rad24/RFC and is critical for full cell cycle arrest, enhancing the activity of Mec1 (Majka et al., 2006). Mec1 phosphorylation targets include Rad53 (ATR in mammalian cells), Rad9 (53BP1 in mammalian cells) and Chk1 (Vialard et al., 1998b) (Chen et al., 2009). Rad9, an essential mediator of cell cycle arrest, which is also phosphorylated by Tel1, binds DNA and stimulates the activation of Rad53 and Chk1 by Mec1 (Vialard et al., 1998b).

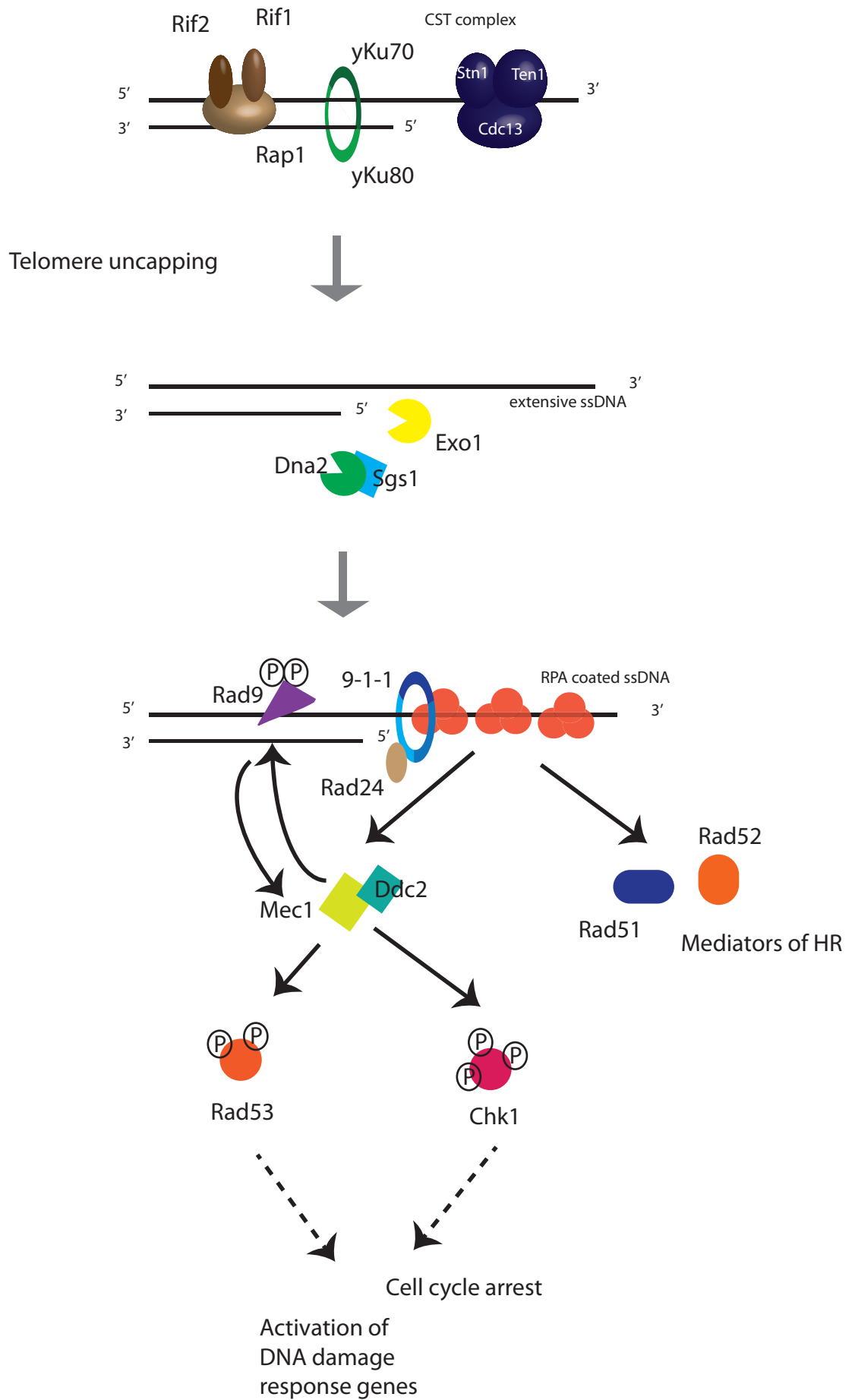


Figure 3 - Telomere uncapping leads to cell cycle arrest. When telomeres become uncapped ssDNA is produced by resection by Exo1 and Dna2/Sgs1. RPA binds to resected ssDNA and activates Rad51 and Rad52, both mediators of Homologous recombination (HR) in addition to Mec1, Mec1 binding partner Ddc2 and the 9-1-1 complex, which is loaded onto DNA by Rad24. Mec1 phosphorylation targets include Rad53, Chk1 and Rad9. Rad9 strengthens the checkpoint response by further activating Mec1, Rad53 and Chk1. Phosphorylation of Chk1 and Rad53 induce cell cycle arrest and activation of DNA damage response genes.



## 1.4 High throughput screens uncover genes involved in telomere biology

Much has been learnt about the cellular responses to telomere uncapping from conditional mutations in telomere capping proteins such as Cdc13 and Stn1 (Addinall et al., 2008) (Holstein et al., 2017). *cdc13-1* is a well studied conditional allele of *CDC13* that causes Rad9 dependant cell cycle arrest at non permissive temperatures due to an accumulation of ssDNA at the telomeres (Garvik et al., 1995). The temperature at which *cdc13-1* can survive (maximum permissive temperature) is therefore increased by deletion of *RAD9* or *EXO1*. High throughput studies using quantitative fitness analysis (QFA) have uncovered many other genes whose deletion suppresses the temperature sensitive phenotype of *cdc13-1* (Addinall et al., 2008). During QFA *cdc13-1* or another gene mutation is cross with a library of gene deletions. Many of these genes identified to increase fitness of *cdc13-1* unsurprisingly function in the DNA damage response but others have less obvious roles in telomere biology, such as the nonsense mediated decay factors (Addinall et al., 2008). High throughput studies yield large amounts of data, however reductionist experiments are often required to understand many of these novel interactions.

## 1.5 Tma20, Tma22 and Tam64

Deletion of *TMA20* and *TMA22* were shown by QFA to strongly suppress the temperature-sensitive phenotype of *cdc13-1* (Addinall et al., 2008). Tma20 and Tma22 are highly conserved genes and the homolog of Tma20 is an oncogene therefore gaining an understanding their role in telomere biology is of relevance to human health. Tma20 and Tma22 were originally identified in yeast as interacting proteins that co-purified with the 40S ribosomal subunit (Fleischer et al., 2006b). The domain organisation of Tma20 and Tma22 is depicted in Figure 4, which illustrates the similarity between yeast, mammalian and Drosophila homologs. The mammalian homolog of Tma20 is Multiple copies in T-cell lymphoma-1 (MCT-1) and the homolog of Tma22 is Density regulated protein 1 (DENR). The N-terminus of Tma20 has an eIF2D\_N domain of unknown function and the C-terminus has a pseudouridine synthase and archaeosine transglycosylase (PUA) domain (Skabkin et al., 2010b). PUA domains have been described as highly conserved RNA binding domains found in a range of proteins in archaea, bacteria and eukaryotes (Cerrudo et al., 2014). Interestingly the PUA domain is also found in Ctf5 and its mammalian homolog

Dyskerin, and mutations in the PUA domain of Dyskerin causes dyskeratosis congenita in humans, a condition characterised by short telomeres (Heiss et al., 1998). The N-terminus of Tma22, and mammalian DENR contains a SUI1/eIF1 domain (also found in eIF1) and the C-terminus contains a SWIB/MDM2 domain (Skabkin et al., 2010b). Tma64 and its mammalian homolog eIF2D (formally called Ligatin) contains in its N terminus eIF2D\_N and PUA of Tma20 and in its C-terminus SUI1/eIF1 and SWIB/MDM2 domains of Tma22. eIF2D has been shown to be functionally homologous to MCT-1 and DENR complex, although deletion of Tma64 does not appear to affect the fitness of *cdc13-1* strains (Skabkin et al., 2010b) (Addinall et al., 2008). The fact that Tma20 contains a PUA domain involved in RNA binding and Tma22 contains a SUI1/eIF1 also found in eIF1 strongly suggests that Tma20 and Tma22 function in translation.

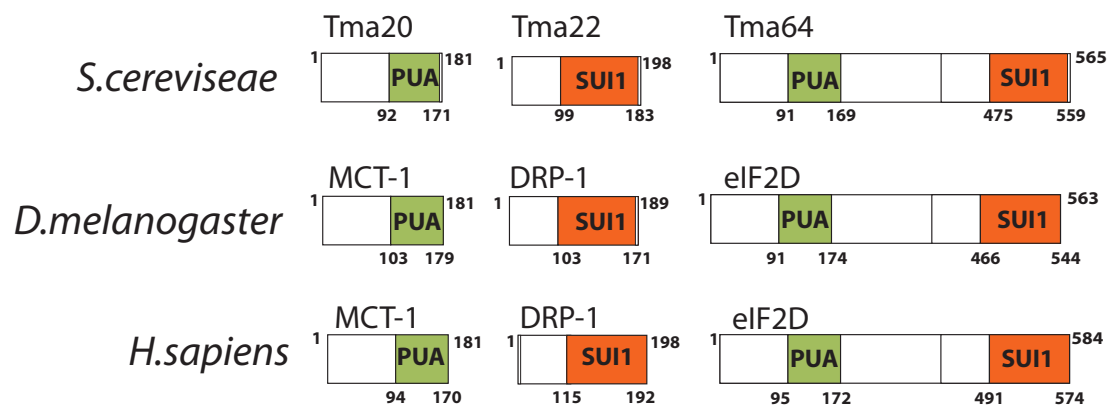


Figure 4: Domain organisation of Tma20, Tma22 and Tma64 in yeast, humans and *Drosophila*. Reproduced from figure in (Skabkin et al., 2010b). Tma20 and homologs in humans and *Drosophila* contain a PseudoUridine synthase and Archaeosine transglycosylase domain (PUA). Tma22 and homologs in humans and *Drosophila* contain a SUI1 domain.

## **1.6 MCT-1 and DENR function in translation**

Recently solved structures of MCT-1 and DENR have shown that they function as a heterodimer (Lomakin et al., 2017). MCT-1 functions in tumorigenesis and various knockdown and overexpression studies have revealed that the disruption of MCT-1 activity affects various cellular functions, including apoptosis, cell cycle progression and DNA damage response (Hsu et al., 2005, Hsu et al., 2007, Herbert et al., 2001). The various phenotypes that have been observed upon overexpression or knockdown of MCT-1 are likely a consequence of altered translation profiles due to the biochemical function of MCT-1 and DENR in translation. Although the role of MCT-1 and DENR in translation is not completely understood current evidence implicate them in and recruitment of Met-tRNA<sub>i</sub><sup>Met</sup> to the initiation complex, recycling of the post termination ribosomes and translation re-initiation. Indeed recent structural studies that confirm their interaction with the 40S ribosomal subunit provide further evidence of their role in translation.

## **1.7 Translation is the synthesis is proteins from mRNAs**

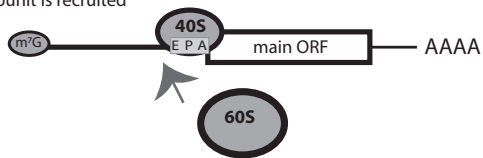
Translation of proteins from mRNA is a tightly regulated process and it has been argued that the dysregulation of translational control can lead to cancer (Ruggero and Pandolfi, 2003). Translation is a cyclical process that involves initiation, elongation, termination and ribosome recycling with translation initiation being the rate limiting stage. Translation initiation involves the assembly of the 80S ribosome onto the initiator methionine. Translation elongation involves the movement of the ribosome along the mRNA, creating a newly synthesised polypeptide chain. An aminoacyl tRNA binds to the A site of a ribosome, and the growing polypeptide chain which is attached to the tRNA positioned in the P site moves to bind to the amino acid attached to the tRNA in the A site. The ribosome then translocates so that the tRNA attached to the polypeptide chain is once again in the P site, leaving the A site empty and the E site occupied by the empty tRNA molecule, where it is subsequently released. Translation termination is the process where a stop codon is encountered by the ribosomal A site and since there is no corresponding aminoacyl tRNA the A site is bound by eukaryotic release factors, which results in the peptide chain being cleaved from the tRNA. During ribosome recycling the tRNA and mRNA are released from the ribosome and the 80S ribosome is dissociated into 40S and 60S subunits.

The 5' cap of mRNA, denoted  $m^7GpppN$ , where N is any nucleotide, is found only on mature mRNAs in which a 7-methylguanosine has been added to the most 5' nucleotide of the transcript leader (TL) by a 5' to 5' triphosphate linkage. The 3' terminus of an mRNA, referred to as the PolyA tail, is comprised of a stretch of adenoside bases and functions in RNA stability.

### Translation initiation

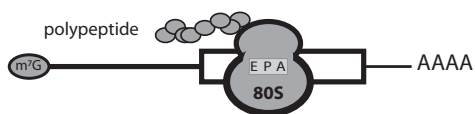
40S ribosome subunit scans the transcript leader until the start codon is located

60S subunit is recruited



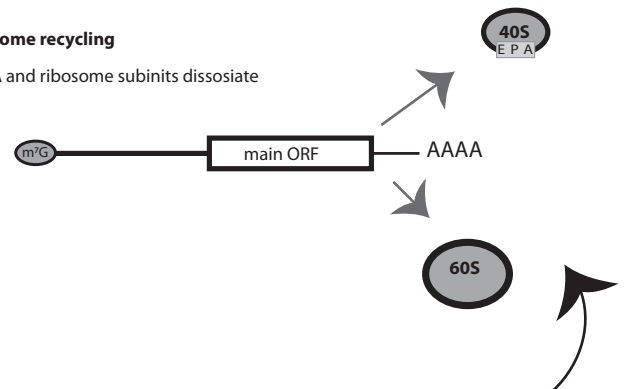
### Translation elongation

Ribosome creates polypeptide chain from mRNA



### Ribosome recycling

mRNA and ribosome subunits dissociate



### Translation termination

Stop codon is encountered polypeptide is released

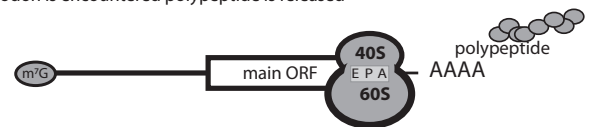


Figure 5. Translation includes initiation, elongation termination and ribosome recycling. During translation initiation the 40S ribosomal subunit scans the TL, in a 5'-3' direction, until it reaches an initiation codon. At this point the 60S joins resulting in an 80S that is poised for translation elongation. During translation elongation the 80S moves along the creating a newly synthesised polypeptide chain. Translation termination occurs when the 80S encounters an in-frame termination codon and results in the release of the polypeptide chain. The mRNA then dissociates from the ribosomal subunits, which split into a 60S and 40S during ribosome recycling.



## 1.8 Translation Initiation

The steps of translation initiation, depicted in Figure 5, are highly conserved in eukaryotes and mediated by eukaryotic translation initiation factors (eIFs). The first step in translation is the formation of the pre-initiation complex (PIC) by the binding of a free 40S ribosomal subunit to the ternary complex (TC), a process facilitated by eIF1, eIF1A and eIF3 (Algire et al., 2002). The TC is comprised of a methionine charged initiator tRNA (Met-tRNA<sub>i</sub><sup>Met</sup>) and a GTP bound eIF2. eIF2 transports Met-tRNA<sub>i</sub><sup>Met</sup> to the AUG of an mRNA and is comprised of 3 subunits,  $\alpha$  (encoded by SUI2),  $\gamma$  (encoded by GCD11) and  $\beta$  (encoded by SUI3). Upon binding of the TC to the 40S, the newly formed complex becomes the 43S PIC complex and is recruited to an mRNA. eIF2 exists in an 'active' GTP bound state or an 'inactive' GDP bound state. eIF2B is the guanine exchange factor (GEF) that displaced GDP and replaces it with GTP (Hinnebusch et al., 2007). eIF2B thus regulates the assembly of active PIC. eIF2B contains 5 subunits which are conserved from yeast to mammals,  $\epsilon$  (encoded by GCD6),  $\delta$  (encoded by GCD2),  $\gamma$  (encoded by GCD1),  $\beta$  (encoded by GCD7) and  $\alpha$  (encoded by GCD3).

mRNA recruitment to 43S PIC is mediated by a protein complex, eIF4F, which binds the m<sup>7</sup>G cap of the TL. Binding of eIF4F to the 43S and attachment to m<sup>7</sup>G cap forms the 48S PIC. eIF4F is comprised of the cap binding factor eIF4E (encoded by CDC33), the scaffold protein eIF4G (encoded by TIF4631) and the RNA helicase eIF4A (encoded by both TIF1 and TIF2). The scaffold protein eIF4G interacts with poly(A)-binding protein (PABP) and promotes binding of eIF4E to m<sup>7</sup>G cap (Tarun Jr and Sachs, 1996) (Ptushkina et al., 1998). eIF4A is a DEAD-box RNA helicase with bi-directional RNA helicase activity and is thought to unwind TL secondary structure (Tanner and Linder, 2001, Rajagopal et al., 2012). The binding of eIF4E to eIF4G is essential for translation initiation and inhibited in mammalian cells by eIF4E binding proteins (4E-BP's). In yeast Caf20 and Eap1 regulate translation by competitively binding to eIF4E (Ptushkina et al., 1998) (Cosentino et al., 2000). In addition to eIF4F, eIF4B and PABP are thought to aid the recruitment of 43S PIC to the 5' end of the TL. eIF4B (encoded by TIF3) binds to 40S ribosomal subunit, inducing a conformational change which is thought to allow the TL of the mRNA to pass through the 43S PIC (Walker et al., 2013).

Upon binding of the 43S PIC to TL m<sup>7</sup>G cap and assembly of the 48S, the ribosome scans until an AUG is within the P site of the ribosome. AUG recognition by the ribosome is aided by the presence of a consensus Kozak sequence (Kozak, 1986a). Although the 43S PIC can bind to 5' cap, scan the TL and, providing eIF1 is present, correctly locate the start codon, helicases eIF4A, eIF4B and eIF4F are required if the TL contains any secondary structure (Pestova and Kolupaeva, 2002). It has however been suggested that, in yeast, Ded and Dpb1 are more important helicases required to resolve TL secondary structure (Berthelot et al., 2004). eIF1, eIF1A, eIF2 and eIF5 are involved in the 'recognition' of the AUG codon, ensuring that the Met-tRNA<sub>i</sub><sup>Met</sup> does not base pair with a non-cognate codon (Hinnebusch et al., 2007). eIF1A is thought to stimulate an open conformation of the ribosome, which returns to a closed conformation when eIF1 is released, following an AUG entering the P site (Maag et al., 2005). Upon codon-anticodon base pairing between the Met-tRNA<sub>i</sub><sup>Met</sup> and the start codon, the GTP bound to eIF2 of the TC is hydrolysed, promoting the dissociation of the TC from the ribosome. TC dissociation from the ribosome allows the binding of a 60S ribosomal subunit, which together with the 40S subunit forms an 80S complex poised for elongation. eIF5 is required for the hydrolyses of GTP by eIF2 which stimulates eIF2 release from the PIC (Jivotovskaya et al., 2006). After joining of the 60S to the 48S PIC a second hydrolysis reaction is required in addition to the release of eIF5 to allow for the 80S to be converted into a state ready for elongation (Hinnebusch et al., 2007). eIF5B, encoded by FUN12, has been demonstrated to catalyse the second hydrolysis reaction (Lee et al., 2002).

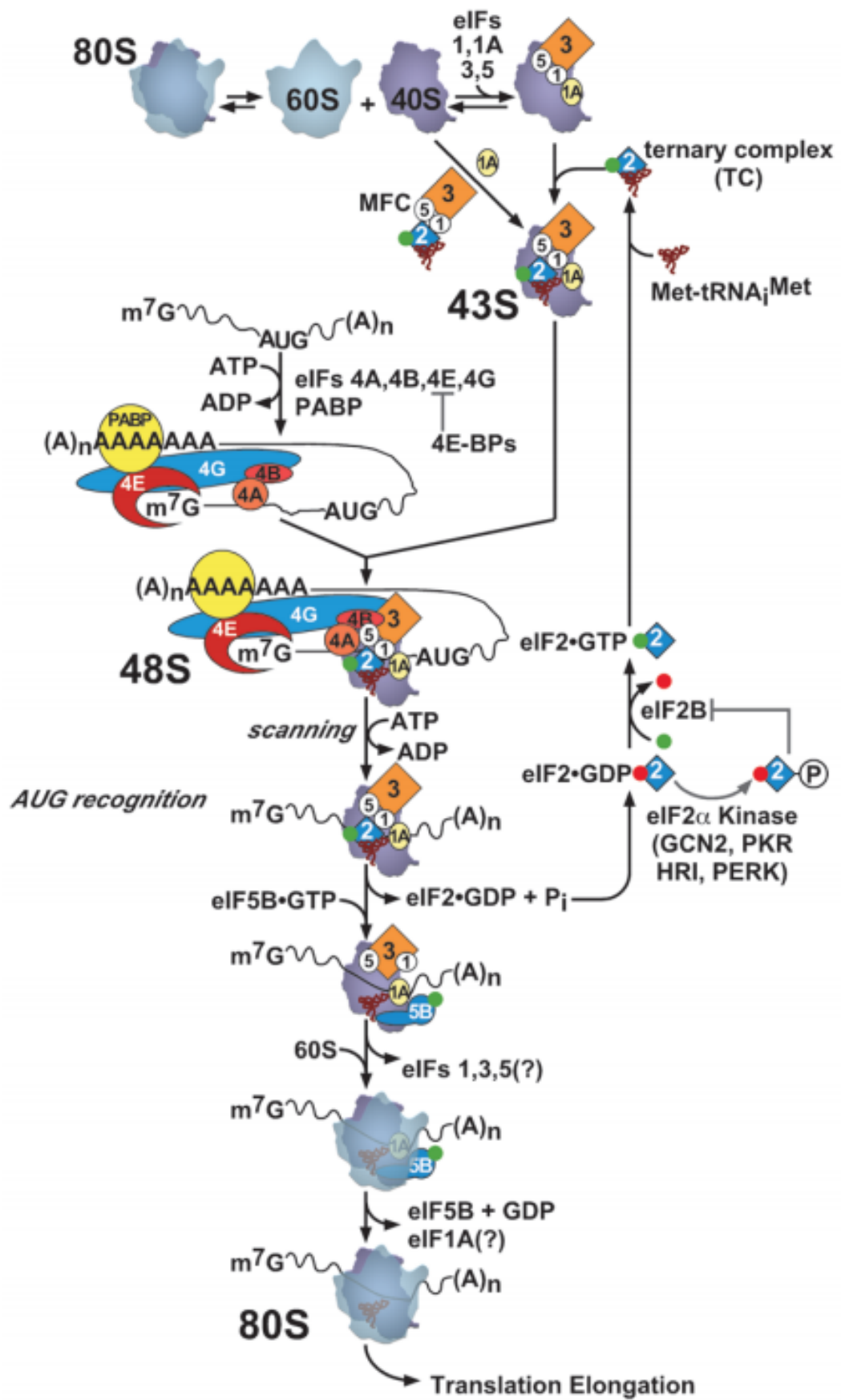


Figure 6: Translation initiation. Taken from (Hinnebusch et al., 2007) The pathway is depicted in a series of discrete steps commencing with the recycling of 80S ribosome into 40S and 60S subunits. The binding of TC to 40S, aided by eIF1A, eIF1, eIF5 and eIF3, forms the 43S PIC. 43S PIC association with the m<sup>7</sup>G cap forms the 48S complex poised for scanning. The ribosome scans the TL in a 5' to 3' direction until an AUG is encountered which triggers GTP hydrolysis of eIF2·GTP and release of eIF2·GDP. eIF2·GDP is converted back into an active TC by GEF eIF2B, a reaction inhibited by phosphorylation of eIF2α. eIF5B promotes binding of the 60S to the 40S to form an elongation poised 80S subunit.

In mammalian cells MCT-1 has been shown to interact with the mRNA 5' cap, in a reaction dependant on eIF4E, and to recruit DENR (Reinert et al., 2006). Luciferase reporter constructs revealed that *in vitro* full length MCT-1, but not MCT-1 lacking the PUA domain, enhance translation, suggesting that MCT-1 promotes translation through binding to RNA via its PUA domain (Reinert et al., 2006). MCT-1 overexpression was shown, by microarray of the mRNAs purified from the polysome fraction, to up-regulate around 25% of the genes which were tested (Reinert et al., 2006).

## **1.9 Some transcripts contain uORFs**

An estimated 50% of human genes and 14% of yeast genes contain upstream open reading frames (uORFs) within their TLs. A uORF is defined as an AUG with an in-frame stop codon within the TL although some definitions include uORFs that have stop codons within the main ORF, termed overlapping ORFs (oORFs) (Figure 7a and b)(Calvo et al., 2009). uORFs generally repress gene expression and this idea is supported by published studies of mRNA and protein abundances in mammalian cells showing that there is a significant decrease in the protein levels of those genes that have uORFs (Calvo et al., 2009). uORFs have also been show to act as negative regulators of translation in yeast, as demonstrated by ribosome profiling experiments (Brar et al., 2012).

### *1.9.1 uORF mediated translational control*

uORFs generally exert negative effect on the expression of the main coding sequence (CDS) since they reduce the number of scanning ribosomes that reach the main initiation codon (Figure 7). This usually occurs as a consequence of the ribosome translating the uORF and dissociating from the mRNA following translation termination at the uORF stop codon (Figure 7c). The ribosome can also stall during elongation of the uORF (Figure 7d). Ribosomal stalling blocks translation of the main ORF by acting as a barrier on the TL, preventing the scanning of other ribosomes (Figure 7d). Further, stalling of the ribosome during elongation or termination of the uORF may induce degradation of the transcript by NMD (Figure 7e). There are two main pathways, which permit the translation of a transcript containing a uORF. The PIC can fail to recognise the uAUG, scan through the uORF and initiate translation at the main ORF, a process known as 'leaky scanning' (Figure 7f). Leaky scanning is

more likely to occur if there is a weak Kozak sequence surrounding the uAUG of the uORF. Secondly following translation the uORF, the 40S ribosome can remain attached to the TL, re-acquire a Met-tRNA<sub>i</sub><sup>Met</sup> and then resume translation at the main ORF in a process known as re-initiation (Figure 7g).

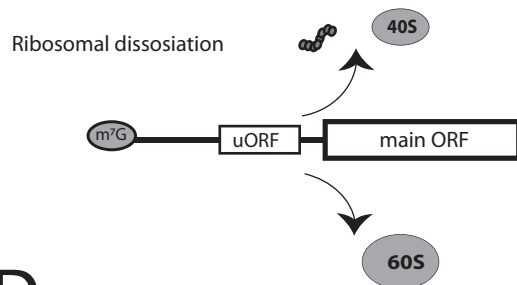
A



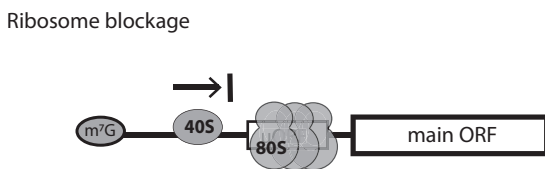
B



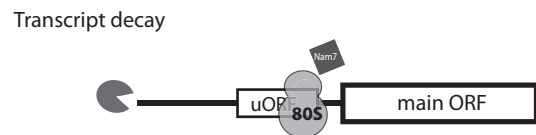
C



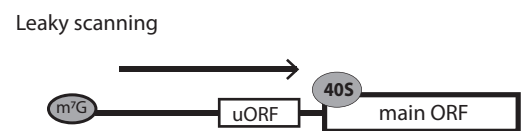
D



E



F



G

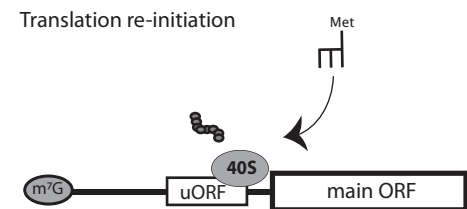


Figure 7: uORF mediated translational control. Reproduced from (Barbosa et al., 2013b) A) Diagram of a transcript with a uORF. B) Diagram of a transcript with an oORF. C) The ribosome dissociates following translation of the uORF. D) The ribosome stalls during translation of the uORF creating a barrier to translation by blocking the progression of other ribosomes. E) The ribosome stalls during translation of the uORF and NMD is activated F) “Leaky scanning” occurs at the uAUG of the uORF and translation of the uORF is bypassed allowing for translation to occur on the main ORF. G) Re-initiation at the main ORF occurs following translation of the uORF.



### *1.9.2 uORFs facilitate expression of some genes in response to stress*

Although uORFs are only found in an estimated 50% of human mRNAs, two-thirds of oncogenes contain uORFs suggesting they may function in ontogenesis (Calvo et al., 2009). This is supported by examples showing that polymorphisms identified in uORFs have been associated with human disease, for example mutations that introduce a uORF into the TL of CDKN2A have been shown to cause familial predisposition to melanoma (Barbosa et al., 2013b) (Bisio et al., 2010). In addition a polymorphism introducing a uORF into the TL of ERCC5 is associated with resistance to platinum based chemotherapy (Somers et al., 2015).

uORFs have been suggested to facilitate the change in expression of a large number of genes in response to changes in cellular activity or stress. Experiments in yeast have indicated that uORFs may change protein expression of a large number of genes to facilitate a switch into meiosis (Brar et al., 2012). uORFs have been suggested to mediate preferential expression of specific mRNAs in response to stress through the integrated stress response (ISR) (Young and Wek, 2016). During the ISR there is a down-regulation in protein production mediated by eIF2 $\alpha$  phosphorylation but curiously a parallel up-regulation in expression of selected genes, which presumably function to ameliorate stress (Young and Wek, 2016). In mammalian cells treated with sodium arsenite there is a global repression of translation but among the 10 genes whose expression was not repressed, 9 of these contain one or more uORFs (Andreev et al., 2015). Further, it has also been suggested that uORFs mediate the preferential translation of some DNA damage response genes when mammalian cells are treated with UVB (Powley et al., 2009). The mechanism that facilitates preferential translation of genes that contain uORFs under conditions of stress is currently unknown. However reductionist experiments have been used to show on specific genes how uORFs facilitate expression in conditions of stress.

### *1.9.3 Specific examples of translation regulation by uORFs*

Delayed translation re-initiation, depicted in Figure 8 has been described where eIF2 phosphorylation results in a change in the start site selection during translation re-initiation. *GCN4* contains 4 uORFs, each within the TL. Following termination and dissociation of 60S subunit and TC after translation of uORF 1 or 2 around 50% of

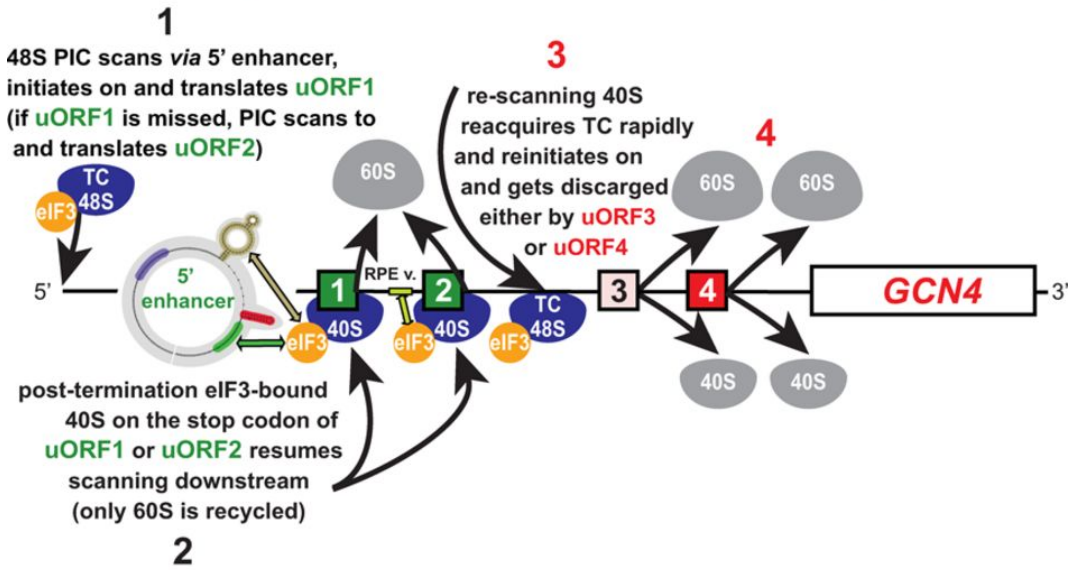
the time the 40S subunit remains attached to the mRNA and continues to scan along the TL (Figure 8) (Gunišová and Valášek, 2014). Under normal conditions 40S subunit re-acquires a TC before reaching the next uAUG and therefore re-initiates translation at uORF3 or 4, subsequently terminating and dissociating from the mRNA without reaching the main initiation codon of *GCN4*. However in conditions eIF2 $\alpha$  phosphorylation there is reduced TC availability meaning the 40S does not re-acquire the TC until reaching the main ORF, thereby bypassing translation of uORF3 and uORF4 and permitting *GCN4* translation (Gunišová and Valášek, 2014). The fact that this model is gene specific is highlighted by the fact that it requires cis acting features within the TL of *GCN4*. Translation re-initiation is permitted following translation of uORF1 and uORF2 by re-initiation permissive elements (RPE) which are positioned within the TL in close proximity to the uORFs (Gunišová and Valášek, 2014). In contrast uORF3 and uORF4 do not efficiently permit translation re-initiation, with uORF4 being considerably less permissive to re-initiation (Gunišová and Valášek, 2014). The delayed translation re-initiation model is conserved and controls expression of the mammalian gene ATF4 in a mechanism analogous to *GCN4*.

eIF2 $\alpha$  phosphorylation has been shown to result in a change in the start site selection causing ribosomes to begin translation at different ATGs. In mammalian cells translation of GADD34 and CHOP are inhibited by uORFs but when eIF2 $\alpha$  becomes phosphorylated bypass of the uORF occurs, allowing initiation at the main coding sequences. The mechanism of different start site selection under conditions of eIF2 however is not yet understood (Young and Wek, 2016).

*CPA1* expression in response to levels of arginine is regulated by a uORF. Under normal conditions the uORF is translated, but translation of *CPA1* is maintained by leaky scanning at the uORF initiation codon (Gaba et al., 2005b). High levels of arginine prevent dissociation of the uORF peptide product, resulting in ribosome stalling (Gaba et al., 2005b). The stalled ribosomes repress translation of the *CPA1* by creating a physically barrier to ribosomes that otherwise would have reached the main ORF by leaky scanning. Additionally the stalled ribosomes further reduce expression since they increase degradation of the transcript by NMD (Gaba et al., 2005b). The amino acid sequence of the uORF is critical to the translational control of *CPA1* and the stop codon has been shown to be dispensable as ribosomes can also stall during elongation (Wang et al., 1998).

Although uORFs generally inhibit downstream translation there are specific examples where uORFs actually promote expression of a corresponding main ORF. For example translation of the uORF of CAT1, a mammalian gene, is thought to unfold repressive TL secondary structure and reveal an IRES, which in turn promotes CAT1 expression (Yaman et al., 2003). Additionally it has been shown that upon exposure to DNA damage, expression of ERCC5 is facilitated by a uORF but inhibited by the same uORF under normal conditions (Somers et al., 2015). The molecular mechanism by which the uORFs promote expression of ERCC5 however is unclear although attributed to the depletion of active eIF2 (Somers et al., 2015).

NON-STARVATION: **GCN4** translation is **REPPRESSED**  
(high levels of TC)



STARVATION: **GCN4** translation is **DEREPRESSED**  
(low levels of TC)

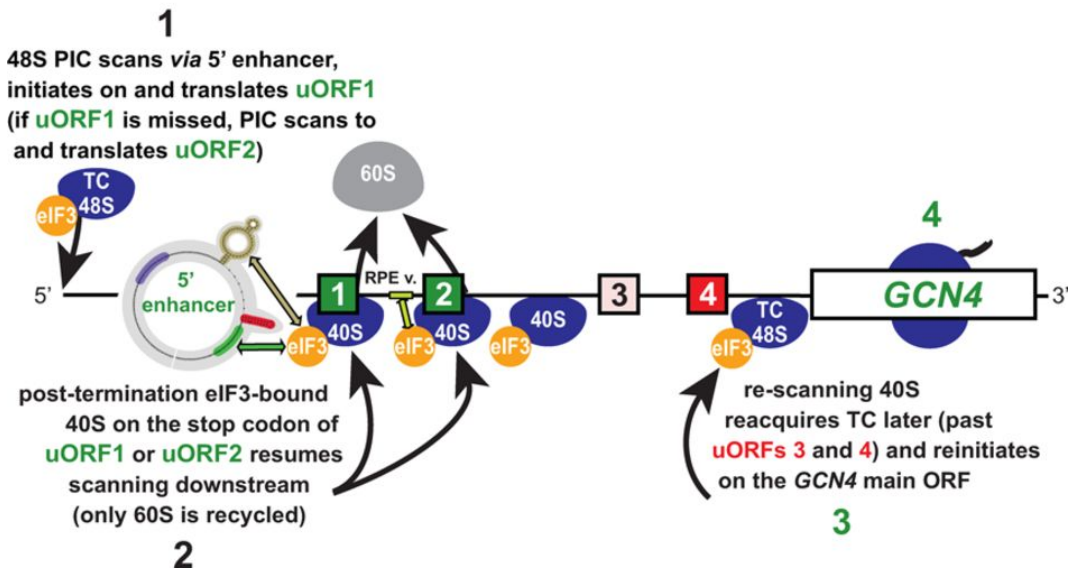


Figure 8. Translation re-initiation controls the expression of GCN4. Figure and legend taken from (Gunišová and Valášek, 2014). Schematic of the *GCN4* TL showing distribution of all four short uORFs, locations of the uORF1-specific and uORF2-specific RPEs, 40S-bound eIF3, and the description of the of the *GCN4* translational control. Upper panel models the events on the *GCN4* TL, which occur under non-starvation conditions with abundant TC levels. The lower panel illustrates the steps that take place under starvation condition with limited supply of the TC.

#### 1.9.4 *Tma20 and Tma22 promote translation of genes that contain uORFs*

The homologs of Tma20 and Tma22, in *Drosophila* were described to regulate the expression of genes that contain uORFs by promoting translation re-initiation (Schleich et al., 2014a). Unlike translation re-initiation described on GCN4 the regulation of re-initiation by DENR is not dependant on regulatory elements within the TL and is not induced in response to low eIF2 $\alpha$  phosphorylation (Schleich et al., 2014a). DENR and MCT-1 are thought to *globally* regulate the translation of genes that contain uORFs whereas the mechanism of translation re-initiation described on GCN4 transcript is unique to GCN4. DENR (Tma22) and MCT-1 (Tma20) were recently also shown to be required for the optimal translation of genes that contain uORFs in human cells although, in contrast with *Drosophila* mammalian homologs of Tma20 and Tma22 only promote translation of transcripts that contain uORFs that are 1 amino acid in length (Schleich et al., 2017). Further the mechanism by which they promote expression after translation of a uORF was not explored, although it could be by translation re-initiation (Schleich et al., 2017). Regulation of translation by DENR (Tma22) and MCT-1 (Tma20) in *Drosophila* and human cells is dependant on the uORF being in a strong initiation context, presumably because the 40S ribosomal complex bypasses uORFs that are in weak initiation contexts (Schleich et al., 2014a) (Schleich et al., 2017).

#### 1.10 Translation termination and ribosome recycling

Translation termination occurs when a stop codon enters the ribosomal A site and is mediated by eukaryotic release factors, eRF1 (encoded by *SUP45*) and eRF3 (encoded by *SUP35*) which form a ternary complex with GTP binding to eRF3. Upon encountering a UAA, UAG or UGA triplet there is no complementary anticodon so the ternary complex is recruited with eRF1 instead entering the A site in place of an acetylated-tRNA (Shoemaker and Green, 2011). Stop codon recognition induces GTP hydrolysis and subsequent eRF3/GDP dissociation (Shoemaker and Green, 2011). eRF1 activates the peptidyl transferase centre of the ribosome which releases the polypeptide by hydrolysing its covalent bond to P-site tRNA (Shoemaker and Green, 2011). The release of polypeptide is promoted by the binding of the recycling factor ABCE1 (Rli1 in yeast) linking termination with recycling (Dever and Green, 2012). ABCE1 mediates the dissociation of 60S from the post termination complex

(post-TC) leaving the 40S subunit bound to mRNA and deacetylated tRNA (Becker et al., 2012) (Pisarev et al., 2010). This reaction is promoted by eIF3 and its loosely associated subunit eIF3j (Hcr1 in yeast) in addition to eIF1 and eIF1A (Pisarev et al., 2007) (Pisarev et al., 2010). The next step involves the release of deacetylated tRNA from the 40S bound to mRNA and is also promoted by eIF1, eIF1A and eIF3 (Pisarev et al., 2007). eIF3j has been suggested to be required for efficient dissociation of mRNA and 40S (Pisarev et al., 2007). Interestingly either eIF2D or MCT-1/DENR have also been shown to substitute for eIF1, eIF1A, and eIF3 to stimulate release of deacylated tRNA and mRNA from the 40S *in vitro* (Skabkin et al., 2010a). Interestingly the NMD has been linked to translation termination and ribosome recycling with *nam7* resulting in an increased stop codon read-through phenotype (Fleischer et al., 2006a) and reduced recycling of post termination ribosomes (Ghosh et al., 2010).

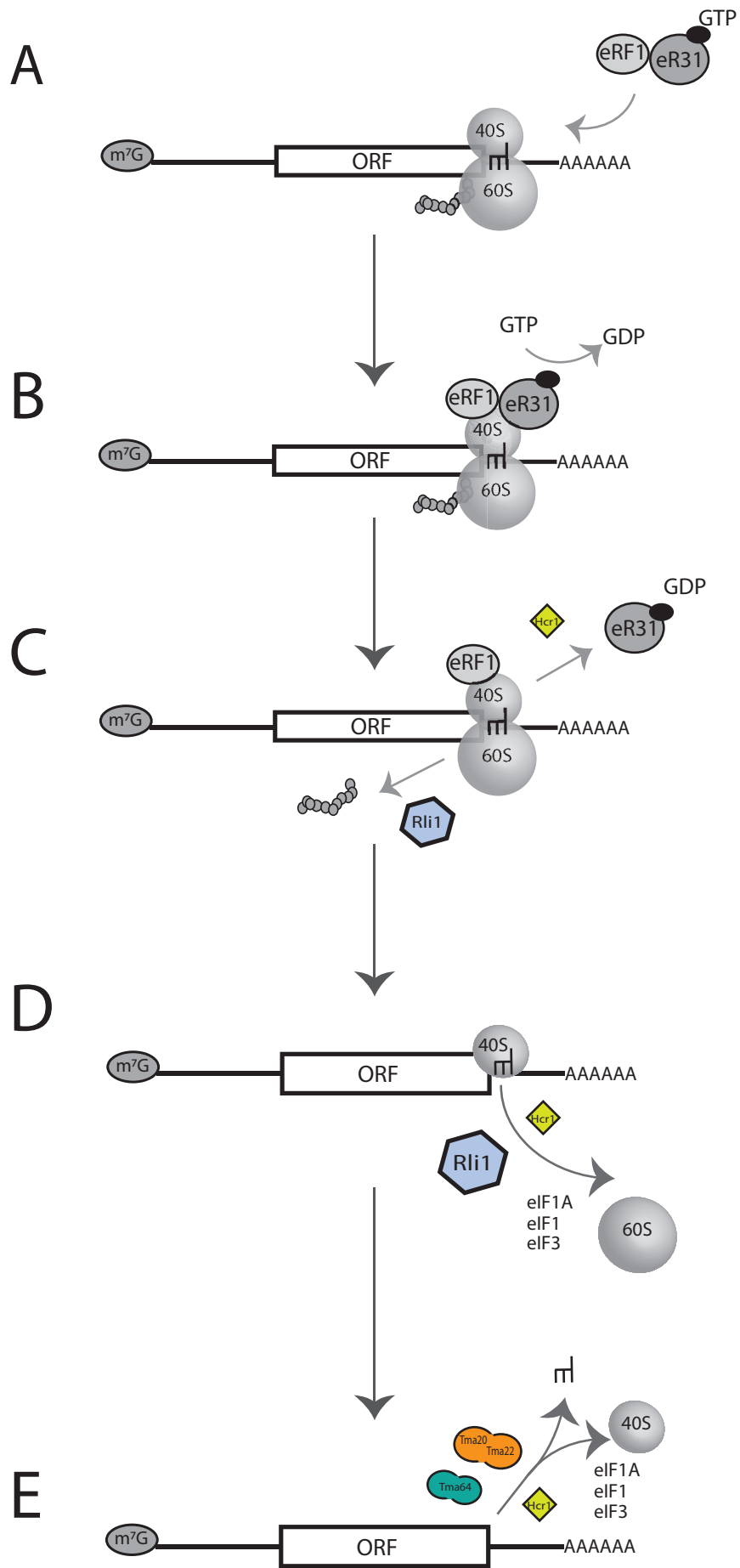




Figure 9. Translation termination and ribosome recycling. Upon stop codon entry into the ribosomal A-site, eRF1 eRF3 are recruited to the ribosome as a ternary complex with GTP. Stop codon recognition induces GTP hydrolysis and eRF3/GDP dissociates. Hcr1 is proposed to promote eRF3·GDP dissociation to allow the recruitment of Rli1. eRF1 then induces release of nascent polypeptide chain and this is also promoted by Rli1. Rli1, Hcr1, eIF1, eIF1A and eIF3 promotes the splitting of the 60S ribosomal subunit. Dissociation of 40S from de-acetylated tRNA and mRNA then occurs, promoted by Hcr1, eIF1, eIF1A and eIF3, Tma20/Tma22 and Tma64.

### 1.11 Translation in conditions of stress

A critical step in translation initiation is the delivery of the Met-tRNA<sub>i</sub><sup>Met</sup> as part of a ternary complex with eIF2 and GTP, to the initiation codon. eIF2 is therefore a critical regulator of translation initiation and depletion of the abundance of active eIF2 (in the GTP bound state) results in a global down regulation in protein synthesis. As a response to a variety of stresses, such as amino acid availability, oxidative stress or heat shock the  $\alpha$  subunit of eIF2 becomes phosphorylated, rendering eIF2 in the inactive state unable to exchange GDP for GTP. The translation of some transcripts can be maintained in low levels of active eIF2 by eIF2 independent recruitment of tRNA<sub>i</sub><sup>Met</sup> to 40S/mRNA binary complex. eIF2 independent translation initiation was originally described on viral transcripts that contain internal ribosome entry sites (IRES) as a mechanism to enable translation while host machinery has been shut down. However since then eIF2 independent translation has been shown in yeast and mammalian cells to allow translation of certain genes in conditions of stress (Gilbert et al., 2007) (Paz et al., 1999).

Evidence suggests that mammalian homologues of MCT-1 (Tma20) and DENR (Tma22) as a complex and eIF2D (Tma64) by itself, promote a type of IRES mediated translation initiation that is independent of eIF2 and also eIF3 (Skabkin et al., 2010b). This mechanism however has only so far been described on viral transcripts where AUG start site must be placed directly in the P-site since. MCT-1 (Tma20) and DENR (Tma22) as a complex or eIF2D (Tma64) alone have more recently shown to promote translation re-initiation on some viral bicistronic transcripts (Zinoviev et al., 2015).

### 1.12 Nonsense mediated Decay

The RNA decay pathway, nonsense mediated decay (NMD) was originally described to target mRNAs which contain premature termination codons (PTC) although the list of NMD targets has since grown to include mRNAs which contain uORFs, long 3'UTRs, frameshifts, unspliced introns and aberrant transcript isoforms generated by different transcription start sites or leaky scanning (Gaba et al., 2005b) (Zaborske et al., 2013) (Celik et al., 2017b). Core NMD factors Upf1 (Nam7), Upf2 (Nmd2) and Upf3 are responsible for targeting an mRNA for NMD with interaction between eRF3 and Nam7 is thought to recruit other factors (Kervestin and Jacobson, 2012). Nam7

is a 5' to 3' RNA helicase central to NMD while Nmd2 and Upf2 are thought to regulate Nam7. Nam7 promotes mRNA decapping via interactions with decapping factors, Edc2, Pat1 and Dcp2 which allows 5'-3' exonucleic degradation performed by Xrn1 (Swisher and Parker, 2011). NMD also stimulates 3'-5' RNA degradation mediated by the exosome and Ski complex and occurs following deadenylation of the polyA tail (Mitchell and Tollervey, 2003). In addition to inducing mRNA degradation which occurs following mRNA uncapping and deadenylation NMD induces translation termination and ribosome dissociation (Kervestin and Jacobson, 2012) .

There is some debate about the mechanism by which translation is terminated and an mRNA is targeted for NMD in yeast and the overlap between these two processes. The core factors of NMD have for example been shown to be involved in normal translation termination.

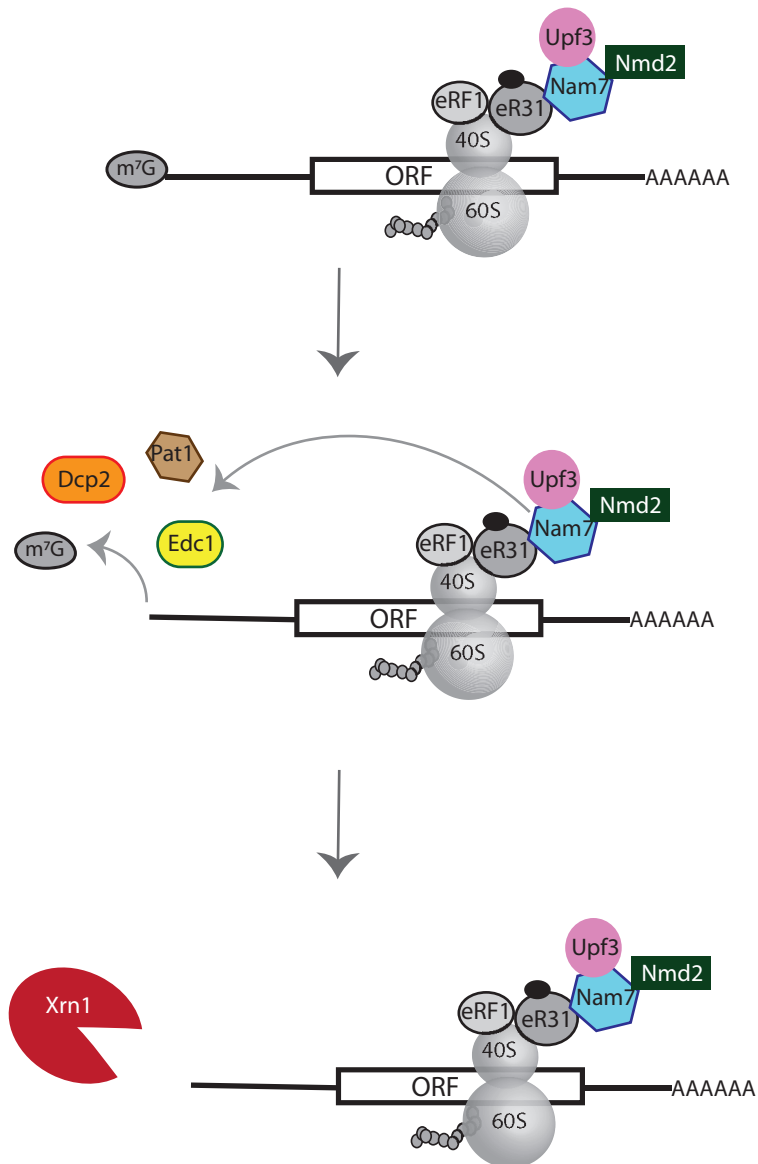


Figure 10: Nonsense mediated mRNA decay. Core NMD factors (Nam7, Nmd2 and Upf3) are recruited to a transcript via interaction of Nam7 with eRF3. This activated decapping factors (Dcp2, Edc1 and Pat1) as well as nucleases 5'-3' and 3'-5' exonucleases.

### 1.13 Aims

The goal of this project is to determine how deletion of *TMA20* and *TMA22* affect the fitness of *cdc13-1* temperature sensitive strains. This will be accomplished by investigating the role of Tma20 and Tma22 in translation and identifying which genes, if any, they regulate the translation of in addition to the mechanism of regulation. An initial aim is to use bioinformatics approaches and published datasets to identify which genes, in yeast, contain uORFs. Molecular genetic techniques will then be used to experimentally test whether candidate genes are under the regulation of Tma20 and Tma22. Yeast genetics will be used to investigate which pathways are affected by the deletion of *TMA20* and *TMA22*.

## **Chapter 2. Materials and Methods**

### **2.1 Recipes for yeast and bacterial media**

#### *2.1.1 Yeast*

Yeast were grown on either YPD, selective dropout media or YDP plus the addition of selective drug. YPD was made using 1% yeast extract, 2% peptone, 2% dextrose and 0.0075% adenine) and selective dropout media 0.13% amino acid dropout powder, 0.17% yeast nitrogen base, 0.5% ammonium sulphate and 2% dextrose. 2% agar was added to make solid plates.

#### *2.1.2 E-coli*

E-coli was grown in Lysogeny Broth (LB) which was made up of 0.5% yeast extract, 1% peptone, 1% NaCL. 2% agar was added to make solid plates. To select cells carrying Ampicillin resistant plasmids 100 mg/ml of Ampicillin was added.

### **2.2 Yeast genetic methods**

#### *2.2.1 Mating*

MATa and MATalpha parental strains were mixed on solid YPD and incubated overnight. Diploids were selected on dropout plates that allow growth of only the diploids.

#### *2.2.2 Sporulation*

Diploid cells were inoculated in 2ml of YPD and grown to saturation on a wheel at 23 °C. 500µL of saturated culture was washed in water and resuspended in ESM media (0.1% yeast extract, 0.05% dextrose, 1% potassium acetate, 0.1% amino acid mix (2g histidine, 10g leucine, 2g uracil). Cultures were incubated on a wheel at 23°C for 3 days or until tetrads were visible.

### 2.2.3 *Tetrad analysis*

Sporulated cultures were washed twice in water and resuspended in 500µL of water. 20µL of culture was incubated at 30°C for 10 minutes with 1.2µL of glucolase (Perkin Elmer) to digest cell walls. Cells were resuspended in 500µL of water and 80µL of tetrads were spread in a line onto YPD plates. Spores were separated using tetrad dissector and allowed to grow for 5 days before being patched out on YPD and replica plated onto appropriate media to determine genotypes.

### 2.2.4 *Gene deletion and epitope tagging*

Gene deletions and epitope tagging were carried out using methods detailed in (Longtine et al., 1998). Selective markers or epitope tags were amplified from plasmids detailed in Appendix B, using primers designed with regions of homology flanking the gene that was to be deleted or flanking the stop codon if an epitope tag was being introduced. PCR product was transformed into appropriate strain using the high efficiency lithium acetate transformation protocol.

### 2.2.5 *High efficiency Lithium acetate transformation*

1ml of stationary phase culture was inoculated in 50ml of YPD with shaking until log phase ( $2 \times 10^7$  cells/mL). Cells were washed twice with water and resuspended in 100 mM Lithium acetate (LiAc). 50µL of the cell suspension was spud down used for transformation. In the following order 240µL 50% PEG, 36µL of 1M LiAc mM, 50µL 2mg/ml salmon sperm DNA and 50µL DNA which is to be transformed were added to the cells and mixed by vortexing. Cells were incubated at 23°C for 30 minutes before heat shock at 42°C for 20 minutes. Cells were resuspended in 200µL of water and plated onto appropriated media at 23°C or 30°C until colonies were visible.

### 2.2.6 *Transformation of plasmids into yeast (One-step transformation)*

Freshly grown cells were re-suspended in 100 µl of one-step buffer (0.2 M Lithium Acetate, 40% PEG, 100 mM DTT). 1µl of plasmid DNA and 5.3 µl of 10 mg/ml salmon sperm DNA was added and mixed by vortexing. Cells were incubated at 45°C for 30 minutes before being plated on appropriate solid selective media.



### *2.2.7 Isolation of plasmids from yeast*

Plasmids were recovered from yeast cells using QIAprep Spin Miniprep Kit, with small adjustments made to the manufacturer's protocol: 50 µl glass beads were added to cells after re-suspension in Buffer P1 and cells were vortexed for 5 minutes. Additionally cells were allowed to incubate at room temperature for 5 minutes after addition of Buffer P2.

### *2.2.8 Growth assays*

2mL cultures were inoculated in YPD overnight on a wheel at 23°C until saturation. The desired dilution series was made in a 96-well plate using water and spotted onto appropriate media using a sterilised replica-plating device.

### *2.2.9 Microscopy*

8 µl of liquid culture was dropped onto a glass slide. Cells were counted/scored and images were photographed under a fluorescence microscope (Nikon eclipse 50i).

## **2.3 Molecular biology methods**

### *2.3.1 Extraction of Genomic DNA from yeast (for PCR)*

When preparing genomic DNA for PCR genomic DNA was extracted from yeast using 'Extraction of Genomic DNA' (Epigenetics) kit according to manufacturer's protocol.

### *2.3.2 Extraction of Genomic DNA from yeast (for Southern Blot)*

For Southern blotting, genomic DNA was prepared using the Yale method. 2 mL saturated cultures, resuspended in 250 µL of buffer (0.1M EDTA (pH7.5), 1:1000 β-mercaptoethanol and 2.5 mg/mL zymolyase 20T (Sigma Aldrich)) and incubated at 37°C for one hour. 55µL of miniprep mix (0.25M EDTA (pH 8.5), 0.5M Tris, 2.5% SDS) was added and suspensions incubated at 65°C for 30 minutes. 68µL of potassium acetate (5M) was added and samples incubated on ice for 30 minutes. Samples were spun down for 20 minutes at 13,000 rpm and the supernatant

transferred to a clean tube. Samples were mixed with 1ml of ethanol and spun down for 10 minutes at 13, 000 rpm for 10 minutes. Pellets were dried and incubated at 37°C for 1 hour in 130µL TE buffer with 1mg/ul RNAase A. 150µL isopropanol was added and samples spun down for 20 minutes at 13, 000 rpm. Pellets were washed with 100µL 70% ethanol. DNA was resuspended in 40 µL TE buffer.

### 2.3.3 Plasmid digestions

Plasmid digestions were set up using 0.5µL of enzyme, 1x digestion buffer. 0.5µL plasmid was used for diagnostic digests and 3µL plasmid was used for cloning reactions. Reactions were incubated at 37° for 2-4 hours.

### 2.3.4 Polymerase chain reaction (PCR)

PCR reactions were set up using the quantities of reagents detailed in the tables below. PCR conditions are also shown. Hot start PCR was used for genotyping whereas ExTaq PCR was used to amplify DNA for cloning, or integrating DNA into the genome.

#### Hot start

F primer (30µM)	0.2µl
R primer (30µM)	0.2µl
MgCl <sub>2</sub> (25mM)	2µl
Hot start polymerase (5u/ul)	0.1µl
dNTPs (2.5mM)	2µl
5x green buffer	4µl
Cells	2µl
water	9.5µl

95°C	15 minutes	1 cycle
94°C	1 minute	30 cycles
55°C	1 minute	
72°C	1 minute per kb	
72°C	10 minutes	1 cycle

## ExTaq

F primer (30μM)	0.4μl
R primer (30μM)	0.4μl
Taq polymerase (5u/ul)	0.2μl
dNTPs (2.5mM)	1.6μl
10x ExTaq Buffer	2μl
DNA (20μg/μl)	2μl
water	13.4μl

95°C	1 minutes	1 cycle
94°C	1 minute	35 cycles
55°C	1 minute	
72°C	1 minute per kb	
72°C	10 minutes	1 cycle

### 2.3.5 Gel electrophoresis

Agarose gels were made up to appropriate concentration with agarose in 0.5x TBE and 1μL per 10μL of SYBR SAFE.

### 2.3.6 In vivo cloning

3μL of plasmid was digested in a 10μL reaction. The insert was amplified using either genomic DNA (extracted using 'Extraction of Genomic DNA from yeast' kit) or plasmid miniprep DNA as template. 10μL vector and 20μL insert were transformed into yeast using 'High efficiency Lithium acetate transformation' protocol. Plasmids were extracted from colonies using 'Isolation of plasmids from yeast' described above.

### 2.3.7 One step Isothermal DNA assembly (Gibson Assembly)

Plasmids were assembled as described in (Gibson et al., 2009). 20μL samples were prepared with 5x isothermal reaction buffer (25% PEG, 500mM Tris-HCl pH 7.5, 50mM MgCl<sub>2</sub>, 50mM DTT, 1mM each dNTPs and 5 mM NAD), 0.8μL 0.2U T5 exonuclease, 4μL 40U taq DNA ligase, 0.5 μL 2U phusion polymerase and 100ng of 6kb fragments of DNA were added in equimolar amounts. Samples were incubated at 50°C for 1 hour and entire reaction was transformed into E-coli.

### 2.3.8 *Plasmid transformation into E-coli*

One Shot® TOP10 Competent Cells (ThermoFisher) were transformed with 1 µL of plasmid DNA as per the manufacturer's instructions, then plated onto LB + Ampicillin and grown overnight hours at 36°C.

### 2.3.9 *Plasmid isolation from E-coli*

Plasmids were isolated from E-coli using QIAprep Spin Miniprep Kit (Quiagen) according to manufacturers protocol.

### 2.3.10 *Luciferase assays*

2ml yeast cultures were grown to exponential phase on a wheel at 30°C and re-suspended in 90µL of Lysis buffer from Dual-Luciferase® Reporter Assay System (Promega). 10µL of cells suspended in Lysis buffer were then added to the well of a 96-well white plate. A PolarStar (Omega) plate reader was used to dispense 50µL of LAR II (Promega), shake for 5 seconds, record the luminescent signal (4 readings with 0.5 second interval times), dispense 50µL Stop & Glo® Reagent (Promega), shake for 5 seconds, and again record the luminescent signal (4 readings with 0.5 second interval times). The mean of the 4 readings that were recorded was used as the final measurement.

### 2.3.11 *TCA protein extraction*

25ml mid log phase cells were washed twice with 2ml of 20% TCA and resuspended in 100µL of 20% trichloroacetic acid (TCA). Cells were frozen for a minimum of 2 hours. Once thawed 50µL glass beads were added to samples and cells lysed using a ribolyser (2x 15s of 6500rpm). 200µL of 5% TCA was added and samples were lysed again using the ribolyser (2x 15s of 6500rpm). Samples were centrifuged at 13,000 rpm for 10 minute and pellets resuspended in 100µL µL Laemmli loading buffer with 5% β-mercaptoethanol (Bio-Rad). 1M Tris was added to neutralise the pH. Samples were boiled for 3 minutes, centrifuged at 13,000rpm for 10 minutes and the supernatant transferred to a clean tube.

### *2.3.12 Western Blotting*

10µL of protein extracts were loaded onto a gradient (4-15%) precast gel (Bio-Rad Mini-Protean TGX) and run for 90 minutes at 100V in Tris/Glycine/SDS running buffer (Bio-Rad). Proteins were blotted, for 30 minutes, onto nitrocellulose membranes using the Trans-Blot® Turbo™ Transfer System (Bio-Rad) according to manufactures protocol. The membrane was washed in 50mL milk (5% (w/v) skimmed milk powder dissolved in PBST) for 30 minutes on a rocking stage at room temperature. The membrane was then incubated in the primary antibody solution (1 in 200 antibody in 1% milk) overnight at 4°C on a rocking stage. The membrane was washed 3 x 10 minutes at the in PBST (PBS with 0.1% Tween20) on a rocking stage before incubation with the secondary antibody solution (1% milk 1 in 2000 HRP conjugated secondary antibody) for 2 hours. The membrane was washed again (3 x 10 minutes in PBST). The protein was detected using Thermo Scientific SuperSignal West Pico Chemiluminescent Substrate according to the manufacturers instructions and imaged on a G box imager (Syngene).

### *2.3.13 RNA isolation*

RNA was isolated from yeast using RNeasy Mini Kit (Qiagen) and RNase-Free DNase Set (Qiagen). 15ml cultures of cells were grown to exponential phase, spun down and washed twice in 1 mL ice-cold DEPC-treated water. The water was removed and cell pellets were incubated at -80°C for at least 2 hours. Pellets were thawed on ice. 100 µL glass beads and 600 µL RLT buffer was added to the samples and cells lysed using a ribolyser (2x 30s of 6500rpm with a 15 second pause). The lysate is transferred to a new eppendorf tube and spud down at 13, 000 rpm for 2 minutes at room temperature. The supernatant was transferred to a clean tube and mixed with 1 volume 70% ethanol. 700µl of the sample was transferred to an RNeasy mini column and spud down for 15 seconds at 12,000rpm. The flow-through was discarded, 350 µl Buffer RW1 was added to the column before being spud down for 15 seconds at 12,000rpm. The flow-through was discarded. 80 µl of DNase I incubation mix (10 µl DNase I stock solution and 70 µl Buffer RDD) was added directly to the column and incubated for 15 minutes at 23°C. Column was spun down with 350 µl Buffer RW1 for 15 seconds at 12,000rpm. 500 µl Buffer RPE was added to the column, spun down for 15 seconds at 12,000rpm then another 500 µl Buffer RPE was added and the column spud down for 2 minutes at 12,000rpm. A final

centrifugation at 13, 000 for 1 minute is carried out with the column in a clean tube. The column is placed into another clean tube, 40µl RNase-free water is added directly to the spin column membrane and spun down for 1 minute at 12, 000 rpm to elute the RNA.

#### *2.3.14 qPCR*

RNA transcript levels are analysed by quantitative reverse-transcriptase PCR (qRT-PCR), using the Superscript III Platinum SYBR green one-step qRT-PCR kit (Invitrogen). 8µl of the reaction mix indicated below was added to 2µl RNA sample (80 ng/µL) in a 96 well plate. An ABI Systems StepOnePlus thermal cycler was used, with the conditions stated in table below.

F primer (10µM)	0.2µl
R primer (10µM)	0.2µl
Superscript III Platinum Taq Mix	0.2µl
2x SYBR Green Reaction Mix	5µl
ROX reference dye	0.2µl
DNA (20µg/µl)	2µl
DEPC-treated water	2.2µl

50°C	3 minutes	1 cycle
95°C	5 minute	1 cycle
95°C	15 seconds	40 cycles
60°C	30 seconds	
40°C	1 minute	1 cycle

### 2.3.15 Southern Blotting

A probe that annealed to the y' elements and TG telomere repeats was constructed using DNA from a plasmid (pDL 987) that contained 120bp of TG repeats and 752bp of the upstream Y' element from telomere VIII-R. The fragment was excised by digestion with BamHI and Xho1 for 3 hours at 37°C (10 µL pDL987 miniprep DNA, 10 units BamHI and 10 units XhoI, 2 µL 10x NEB Cutsmart Buffer, 7µL sterile water). Digested fragment was excised from a gel and purified using a Qiaprep Gel Extraction Kit (Qiagen) according to manufacturers protocol. The probe was Digoxigenin-dUTP (DIG) labelled using random oligonucleotide primers following manufacturers guidelines in the DIG High Prime Labelling and Detection Starter Kit II (Roche).

DNA was extracted as described in 2.3.2 and samples were equalised to the same concentration by running 2 µL on a 1% gel and measuring band intensity using GeneSys (Syngene). 2.5 µL of equalised DNA was run on a 1% agarose gel at 22 volts for 16 hours. The gel was imaged on a G-box (Syngene) to indicate the relative amount of DNA in each sample. The gel was incubated in 0.25M hydrochloric acid for 15 minutes, rinsed twice in water and incubated for 30 minutes in 0.5M sodium hydroxide for 30 minutes. The gel was blotted to a nylon membrane (Roche) using a vacuum blotter (Model 785, BioRad) at 5 inches Hg in 10x saline-sodium citrate

(SSC; 0.15M sodium citrate; 1.5M NaCl) for 90 minutes. The DNA was cross-linked to the membrane using auto UV-crosslinking (Stratalinker), rinsed in water and then air-dried for 30 minutes. The probe was hybridized according to manufacturers instructions using the labelled probe and DIG High Prime Labelling and Detection Starter Kit II (Roche). The membrane was imaged using G-box (Syngene).

## **2.4 Bioinformatics**

### *2.4.1 Obtaining TL sequences*

Gene specific TL lengths were obtained from Arribere *et al* (2013) and Nagalakshmi *et al* (2008). Sequences were retrieved using SGD API.

### *2.4.2 Scoring initiation contexts*

A Kozak matrix was constructed using frequency of occurrence of each nucleotide in positions -1, -2, -3, -4, +1 and +2 relative to the ATG of the main translation start site of each gene in the yeast genome. Sequences were downloaded from Biomart (<http://www.biomart.org>; Ensembl Genes 79, Accessed: Feb 2015). The Kozak Score of a given uORF was generated using a multiplicative model whereby the 'frequency of occurrence of nucleotides' in Figure 20a, from position minus 4 to plus 2 relative to the ATG were multiplied together. For example the consensus Kozak sequence of AAAAatgAC would yield a Kozak Score of  $[0.42902 \times 0.576061 \times 0.397639 \times 0.439928 \times 0.317693 \times 0.366557]$  which is 0.005035.

### *2.4.3 Analysis of TLs*

The uAUGs in every TL was identified along with a Kozak score describing its initiation context, whether it was part of a uORF, the length and position of uORF from the main ORF. The reading frame of the uORF was also identified.



## Chapter 3. *TMA20* and *TMA22* act in the same pathway as NMD genes to increase fitness of *cdc13-1* cells

### 3.1 *Tma20* and *Tma22* act as interacting partners to decrease fitness of *cdc13-1* and *stn1-13*.

*TMA20* and *TMA22* are two genes identified by QFA to be very strong suppressors, when deleted, of the temperature sensitive growth defect of *cdc13-1* (Addinall et al., 2011). To confirm that deletion of *TMA20* and *TMA22* increase fitness of *cdc13-1*, *tma20Δ* and *tma22Δ* were introduced into an independent genetic background (W303) and fitness of *cdc13-1* was compared with fitness of *cdc13-1 tma20Δ* and *cdc13-1 tma22Δ* using a growth assay (spot test). A moderate increase in growth was observed when *tma20Δ* or *tma22Δ* were combined with *cdc13-1* and grown at semi permissive temperatures of 27°C and 28°C (Figure 11a). Compared with some gene deletions, such as *RAD24*, which permit growth of *cdc13-1* at 29°C *TMA20* and *TMA22* can be described as moderate suppressors of *cdc13-1* thermo-sensitivity (Addinall et al., 2011). At temperatures above 29°C *cdc13-1 tma20Δ* and *cdc13-1 tma22Δ* growth resembles that of *cdc13-1* indicating that *Tma20* and *Tma22* do not contribute to the decrease in fitness of *cdc13-1* mutation at higher temperatures (Figure 11a). *Tma64* resembles a fusion of *Tma20* and *Tma22* where the N-terminus of *Tma64* shares homology with *Tma20* and the C-terminus shares homology with *Tma22* (Figure 4). Additionally *Tma20* and *Tma22* have been suggested to be functional homologues of *Tma64* (Skabkin et al., 2010a). We therefore hypothesised that *Tma64* may play a similar role to *Tma20* and *Tma22* at telomeres and to test this *TMA64* was deleted in a *cdc13-1* background. Interestingly *tma64Δ* has no effect on the fitness of *cdc13-1*, as demonstrated by the comparable fitness of *cdc13-1* and *cdc13-1 tma64Δ*, suggesting that in yeast *Tma64* has a distinct function from *Tma20/Tma22* complex (Figure 11a). It has been previously demonstrated that *Tma20* and *Tma22* physically interact and knockdown of the mammalian and *Drosophila* homolog of *Tma20*, MCT-1, both lead to a drastic reduction in DENR and vice versa (Fleischer et al., 2006a) (Skabkin et al., 2010a, Schleich et al., 2017). To explore if there is genetic redundancy between *TMA20* and *TMA22* we combined deletion of both genes in *cdc13-1* background and observed no difference in the fitness of *cdc13-1 tma20Δ tma22Δ* compared with *cdc13-1 tma20Δ* and *cdc13-1*

*tma22Δ*. This supports a model that Tma20 and Tma22 function as a complex in yeast (Figure 11a).

To further explore the role of Tma20 and Tma22 at telomeres we tested if they also contribute to the loss of fitness of cells harbouring different telomere defects. We used *stn1-13*, a temperature sensitive allele of *STN1* which results in loss of fitness at temperatures higher than 30°C. Although Stn1 and Cdc13 are both members of the CST complex it is thought that they also perform distinct functions, therefore it is possible that Tma20 and Tma22 may specifically affect cells with defective Cdc13 but not defective Stn1 (Holstein et al., 2014). *tma20Δ* and *tma22Δ* were combined with *stn1-13* and fitness of cells was assessed using a spot test (Figure 11b). We found that while *stn1-13* grew worse than wild type at 32°C - 37°C, growth of *stn1-13 tma20Δ*, *stn1-13 tma22Δ* and *stn1-13 tma20Δ tma22Δ* were moderately improved (Figure 11b). This indicates that Tma20 and Tma22 contribute to the lethality of cells with defective CST via a function that is common to both Cdc13 and Stn1, most likely telomere capping rather than for example regulating telomerase. Telomerase recruitment is promoted by Cdc13 and inhibited by Stn1 (Puglisi et al., 2008, Nugent et al., 1996b).

To examine if Tma20 and Tma22 specifically affect fitness of cells with telomere damage in the form of defective CST complex we measured the effect of *tma20Δ* on fitness of *yku70Δ*. yKu70 and yKu80 form a heterodimer essential to NHEJ DNA repair pathway. yKu70/yKu80 also bind the telomeres, and function in chromosome end protection, (Maringele and Lydall, 2002a) telomerase recruitment (Fisher et al., 2004, Stellwagen et al., 2003) and silencing of nearby genes (Boulton and Jackson, 1998). Strains carrying *yku70Δ* exhibit a temperature sensitive growth defect above 36°C, which is attributed to telomere defect (Maringele and Lydall, 2002a). In contrast with the effect of *TMA20* deletion on fitness of cells with *cdc13-1* and *stn1-13*, *tma20Δ* has no effect on the fitness of *yku70Δ* since growth of *yku70Δ* resembled *yku70Δ tma20Δ* at all temperatures tested (Figure 11c). We can therefore conclude that Tma20 and Tma22 contribute to the loss of fitness of cells lacking a functional CST complex but do not affect fitness of cells with all telomere defects, such as *yku70Δ*.

Although ssDNA accumulation occurs as a result of telomere uncapping in both *yku70Δ* and cells with defective CST (*cdc13-1*), cells respond slightly differently to

damage induced by *yku70Δ* and *cdc13-1* (Maringele and Lydall, 2002a). The 9-1-1 complex (Rad17, Ddc1 and Mec3) and Rad24, which loads the 9-1-1 complex onto damaged DNA, is required for cell cycle arrest in *cdc13-1* but not *yku70Δ* (Maringele and Lydall, 2002a). This suggests that Tma20 and Tma22 may modulate levels of some DNA damage response proteins, such as Rad17, Ddc1, Mec3 or Rad24 that respond to *cdc13-1* and not *yku70Δ*.

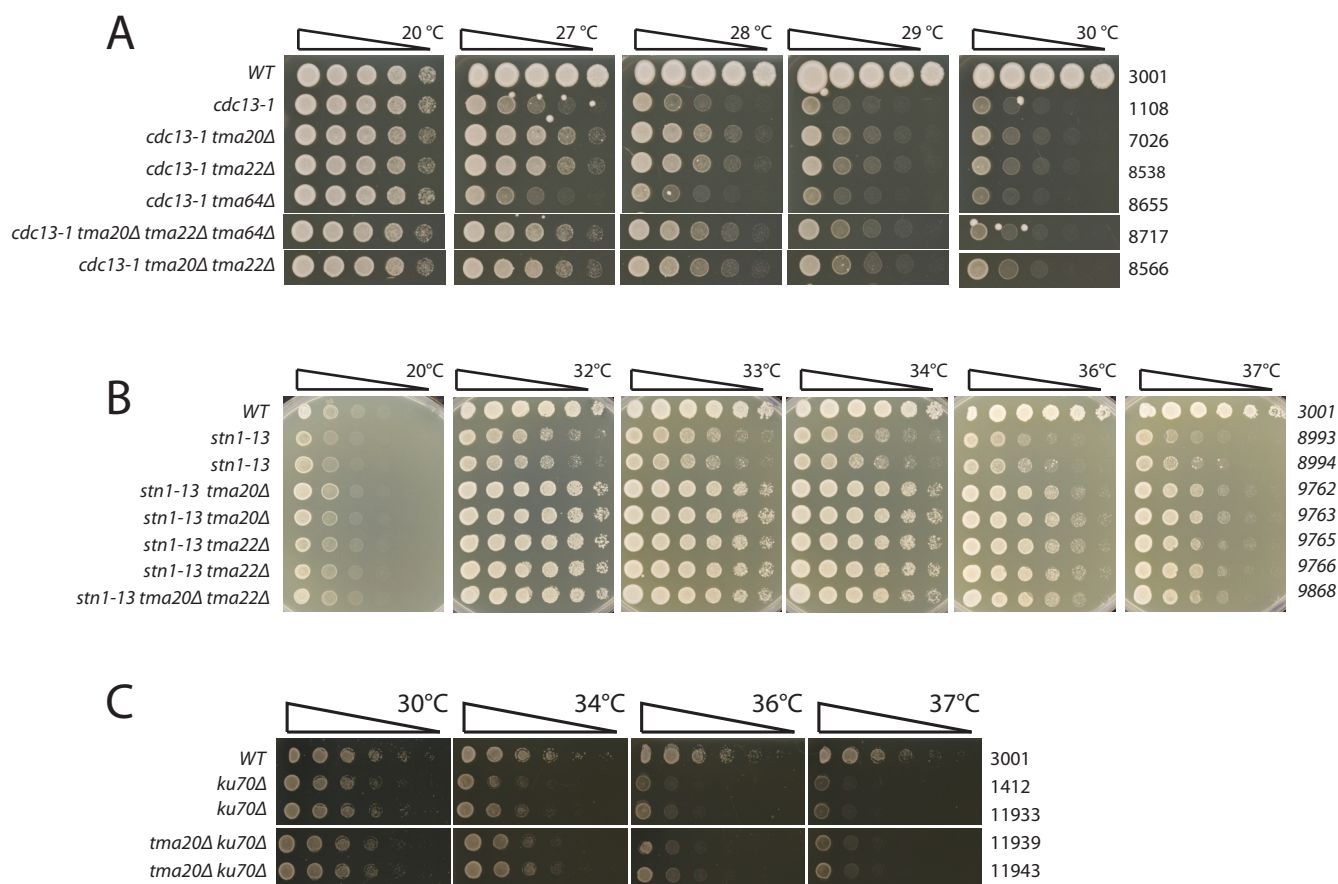


Figure 11: Tma20 and Tma22 act as interacting partners to decrease fitness of *cdc13-1* and *stn1-13*. Yeast strains of the indicated genotypes were grown to saturation in liquid YEPD, serially diluted, 5-fold, in water, and spotted onto YPD solid media. Plates were incubated for two days at the indicated temperatures before being photographed. Genotypes and strain numbers are listed on either side of the spot tests.

### 3.2 Tma20 does not affect telomere length

Some gene deletions that improve fitness of *cdc13-1*, such as deletions in the NMD pathway also affect telomere length and mutations that allow growth in the absence of *CDC13* have been shown do so by rearranging their telomeres (Larrivee and Wellinger, 2006). To test whether deletion of *TMA20* and *TMA22* affect the integrity of the telomere DNA telomeres were examined using Southern blot with a Y' and TG repeat probe to detect terminal y' elements (probe is indicated in Figure 12a). For comparison we also measured the length of telomeres in *nmd2Δ*.

Consistent with previous observations we saw that telomeres in *nmd2Δ* were shorter than telomeres in WT strains (Dahlseid et al., 2003). In contrast with this *tma20Δ* has no effect on telomere length (Figure 12b). Telomeres of *tma20Δ* appear completely normal and indistinguishable from telomeres in WT cells (Figure 12b). The observation that *tma20Δ* does not induce any telomere re-arrangements is in line with the observation that *tma20Δ* is a weak suppressor of *cdc13-1*. Further the lack of phenotype observed in the telomeres of *tma20Δ* supports the idea that the function of Tma20 at telomeres is indirect.

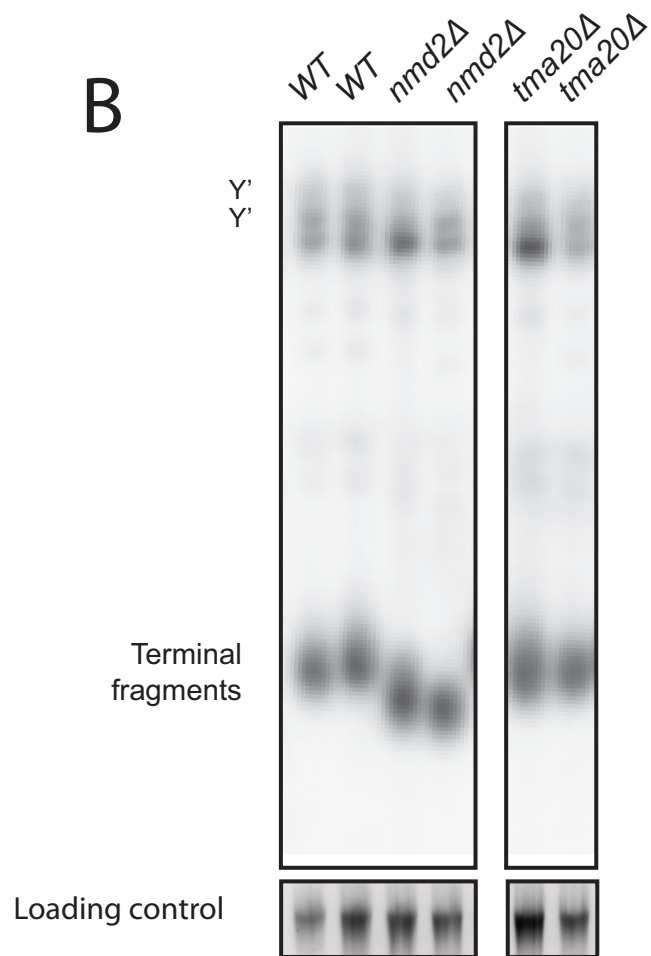
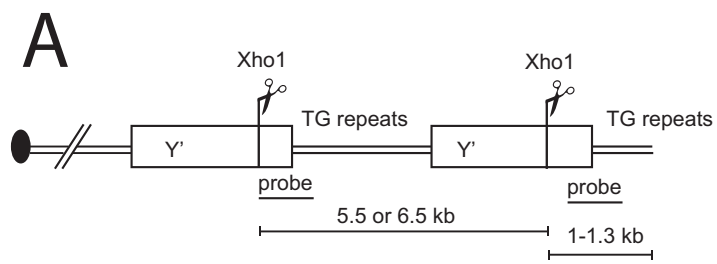


Figure 12. Tma20 does not affect telomere length. A) Diagram indicating the positions of Xho1 sites where the Y' and TG1-3 repeat probe binds. B) Genomic DNA was isolated from yeast strains that are indicated. Telomere structures were analyzed by Southern blotting (see methods) using the Y' and TG probe indicated. Digested DNA stained with SyBr Green was used as a loading control.



### 3.3 *HCR1* strongly reduces fitness of *cdc13-1* cells and acts in a different pathway as *TMA20* and *TMA22*.

Since Tma20 homologs in mammalian cells have been implicated in ribosome recycling (illustrated in Figure 9) and translation re-initiation we decided to explore the possibility that other genes proposed to be involved in recycling and translation re-initiation may also affect fitness of *cdc13-1*. *HCR1* is a non-essential gene suggested to promote ribosome recycling and forms a complex with eIF3, an initiation factor thought to facilitate translation re-initiation (Beznoskova et al., 2013) (Munzarova et al., 2011) (Roy et al., 2010). Although the role of Hcr1 in ribosome recycling in yeast is unclear the mammalian homolog, eIF3j, has been proposed to promote release of 60S as well as subsequent dissociation of 40S with mRNA and tRNA (Figure 9) (Pisarev et al., 2007) (Pisarev et al., 2010). *hcr1Δ* and *tma20Δ* have also both been reported to increase stop codon read-through, further supporting the hypothesis that they may have similar roles (Fleischer et al., 2006a, Beznoskova et al., 2013).

To examine if *HCR1* also affects fitness of cells with *cdc13-1* telomere defect *hcr1Δ* was combined with *cdc13-1* and growth of resulting strains were measured by spot test. A moderate increase in growth is observed in *cdc13-1 hcr1Δ* compared with *cdc13-1* at 26°C, 27°C, 28°C and there is even a slight increase in growth observed at 29°C (Figure 13). Interestingly at 28°C and 29°C *cdc13-1 hcr1Δ* grows better than *cdc13-1 tma20Δ* demonstrating that *HCR1* is a stronger suppressor of *cdc13-1* temperature sensitivity than *TMA20* (Figure 13). In order test whether *HCR1* acts independently of *TMA20* and *TMA22* to suppress *cdc13-1* we combined *hcr1Δ* with *cdc13-1 tma20Δ* and *cdc13-1 tma22Δ* and compared growth of resulting strains by spot test. At 27°C-29°C *hcr1Δ* improves the growth of *cdc13-1 tma20Δ* and *cdc13-1 tma22Δ* suggesting that *TMA20* and *TMA22* act in a different pathway to *HCR1* (Figure 13). At all temperatures fitness of *cdc13-1 hcr1Δ tma20Δ tma22Δ* was comparable to fitness of *cdc13-1 hcr1Δ tma20Δ* and *cdc13-1 hcr1Δ tma22Δ* consistent with a model of Tma20 and Tma22 being interacting partners (Figure 13).

Our data indicate general defects in ribosome recycling may facilitate growth of cells with *cdc13-1* since deletion of either *HCR1* or *TMA20/TMA22* both improve fitness of *cdc13-1*. We also found that *HCR1* and *TMA20/TMA22* act independently to affect fitness of *cdc13-1* supporting the idea that they have distinct functions in translation.

It is unclear if Hcr1 or Tma20/Tma22 impact fitness of *cdc13-1* by affecting the efficiency of ribosome recycling or through a different mechanism such as translation initiation.

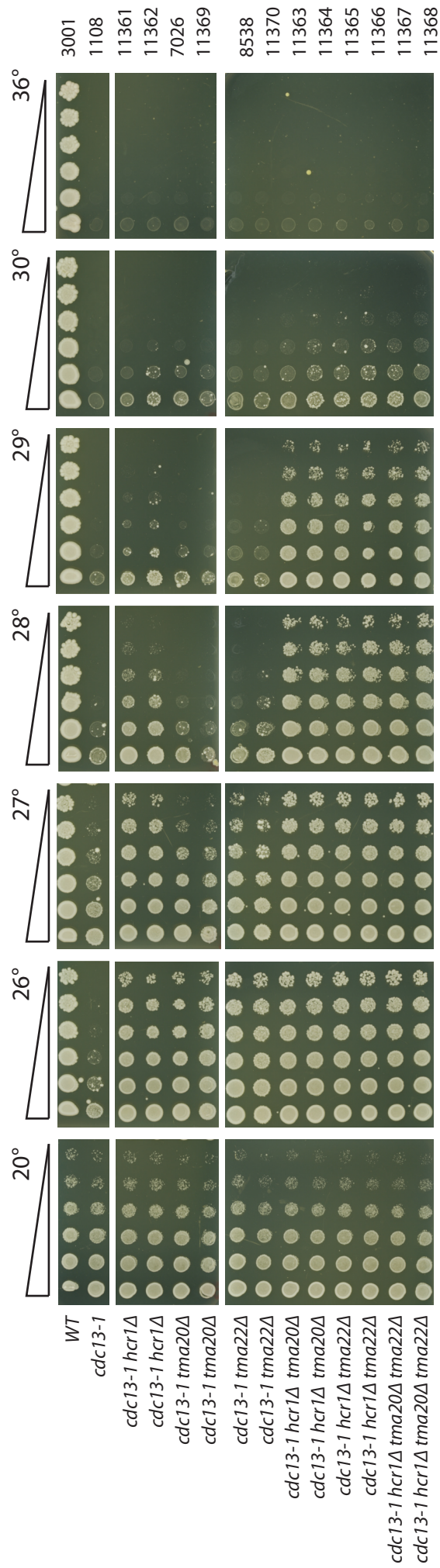


Figure 13. *HCR1* strongly improves fitness of *cdc13-1* cells and acts in a different pathway as *TMA20* and *TMA22*. Yeast strains of the indicated genotypes were grown to saturation in liquid YEPD, serially diluted, 5-fold, in water, and spotted onto YPD solid media. Plates were incubated for two days at the indicated temperatures before being photographed. Genotypes and strain numbers are listed on either side of the spot tests

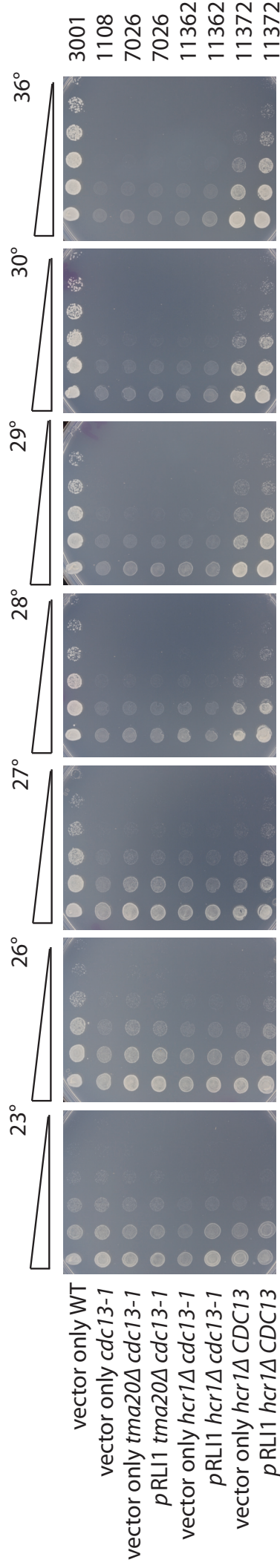
### 3.4 *cdc13-1 hcr1*Δ, but not *cdc13-1 tma20*Δ, is suppressed by overexpression of *RLI1*

To further explore the relationship between a possible role of Tma20/Tma22 or Hcr1 in ribosome recycling and their function in decreasing fitness of *cdc13-1* we decided to test if *cdc13-1 hcr1*Δ or *cdc13-1 tma20*Δ was suppressed by overexpression of the ribosome-recycling factor Rli1. Rli1 is required for efficient ribosome recycling, promoting 60S subunit dissociation (Shoemaker and Green, 2011). Overexpression of Rli1 in yeast has been described to suppress the slow growth and increased stop codon read-through phenotype of *hcr1*Δ further supporting the idea that *hcr1*Δ promotes ribosome recycling (Beznoskova et al., 2013). We speculated that overexpression of Rli1 would also suppress ability of *hcr1*Δ to increase fitness of *cdc13-1*. The observation that mammalian homologs of Tma20 and Tma22 also promote ribosome recycling made us hypothesise that overexpression of Rli1 may decrease the fitness of *cdc13-1 tma20*Δ by restoring efficient ribosome recycling.

To explore the role of ribosome recycling in contributing to the lethality of *cdc13-1*, Rli1 was cloned into a 2μ expression vector and transformed into *hcr1*Δ, *cdc13-1 hcr1*Δ and *cdc13-1 tma20*Δ. Consistent with published observations the growth rate of *hcr1*Δ is improved by overexpression of Rli1- indicated by better growth after 1 day, but equal growth after 3 days, of *hcr1*Δ *RLI1* compared with *hcr1*Δ, between 27°C and 36°C (Figure 14a and b). *hcr1*Δ and *hcr1*Δ *RLI1* were comparable at 23°C, presumably because all strains grow slowly at low temperatures. After 1 day, growth of all strains harbouring *cdc13-1* allele was similar, however differences in growth could be observed after 3 days (Figure 14 and b). Growth of *cdc13-1* and *cdc13-1 pRLI1* was similar at all temperatures tested showing that overexpression of ribosome recycling factor, Rli1 does improve fitness in *cdc13-1* which indicates *cdc13-1* are not defective in ribosome recycling. Interestingly however, while *cdc13-1 hcr1*Δ grew considerably better than *cdc13-1* at 26°C- 29°C, growth of *cdc13-1 hcr1*Δ transformed with *pRLI1* and *cdc13-1* are similar showing that Rli1 overexpression decreases fitness of *cdc13-1 hcr1*Δ (Figure 14b). At 26°C- 28°C *cdc13-1 tma20*Δ grew better than *cdc13-1* but similar to *cdc13-1 tma20*Δ *pRLI1* showing that, in contrast with the effect of Rli1 overexpression in *cdc13-1 hcr1*Δ, overexpression of Rli1 does not suppress *tma20*Δ (Figure 14b).

The finding that *cdc13-1 hcr1*Δ is suppressed by overexpression of Rli1 supports the idea that Hcr1 decreases fitness of *cdc13-1* by promoting ribosome recycling. Rli1 promotes peptide release and 60S dissociation (Shoemaker and Green, 2011) suggesting the increase in fitness of *cdc13-1* that occurs upon deletion of *HCR1* is a consequence of defective peptide release or 60S dissociation induced by *hcr1*Δ. It is also possible that the decrease in growth rate that occurs upon deletion of *HCR1* improves fitness of *cdc13-1*. In contrast Rli1 does not affect fitness of *cdc13-1 tma20*Δ, which is consistent with the idea that Tma20 promotes subsequent dissociation of tRNA/mRNA/40S (Skabkin et al., 2010a). It is unclear if the increase in fitness of *cdc13-1* by *tma20*Δ is a result of defective dissociation of tRNA/mRNA/40S.

# A



# B

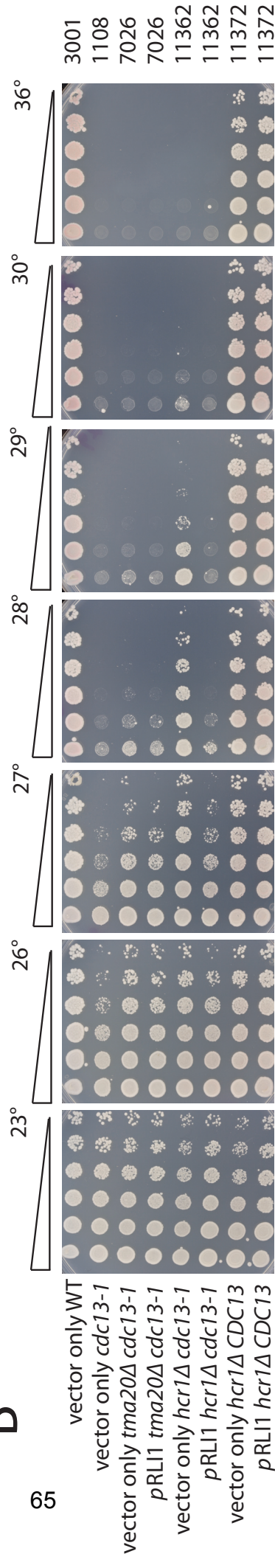


Figure 14: *cdc13-1 hcr1* $\Delta$ , but not *cdc13-1 tma20* $\Delta$ , is suppressed by overexpression of *RLI1*. *RLI1* was cloned into a high copy vector and transformed into the indicated yeast strains. Cells were grown to saturation in media lacking leucine, serially diluted, 5-fold, in water, and spotted onto solid media lacking leucine. Plates were incubated for two days at the indicated temperatures before being photographed. Genotypes and strain numbers are listed on either side of the spot tests.



### 3.5 *TMA20* and *TMA22* act in the same pathway as *NMD* to affect fitness of *cdc13-1*

Since Tma20/Tma22 were shown to act independently of recycling factors Hcr1 and Rli1 to suppress *cdc13-1* we wondered if they function in the same pathway as other known suppressors of *cdc13-1* temperature sensitivity that also associate with translation termination. Deletion of Nonsense mediated decay (NMD) genes *NAM7*, *NMD2*, *UPF3* and *EBS1* strongly increase fitness of *cdc13-1* (Addinall et al., 2011) and have also been shown to associate with the translation machinery and to promote ribosome recycling and increase stop codon read-through (Fleischer et al., 2006a) (Ghosh et al., 2010).

To investigate if deletion of *TMA20* and *TMA22* may therefore suppress the temperature sensitive phenotype of *cdc13-1* strains through the same pathway as NMD deletions *nmd2Δ* and *ebs1Δ* were combined with *tma20Δ* and *tma22Δ* in a *cdc13-1* background and fitness assessed by spot test. There was no observable difference in growth of *cdc13-1 nmd2Δ tma20Δ*, *cdc13-1 nmd2Δ tma22Δ* and *cdc13-1 nmd2Δ* in any of the temperatures that were tested indicating that *TMA20* and *TMA22* act in the same pathway as *NMD2* to suppress *cdc13-1* (Figure 15). *cdc13-1 ebs1Δ tma20Δ* and *cdc13-1 ebs1Δ tma22Δ* grew slightly better than *cdc13-1 ebs1Δ* at 28°C however neither of *ebs1Δ tma20Δ*, *ebs1Δ tma22Δ* or *ebs1Δ* suppressed the temperature sensitive phenotype of *cdc13-1* at 29°C or above (Figure 15). The fact that the increased fitness observed in *cdc13-1 ebs1Δ tma20Δ* and *cdc13-1 ebs1Δ tma22Δ* compared with *cdc13-1 ebs1Δ* was only small suggests that Ebs1 and Tma20 or Ebs1 and Tma22 are not epistatic. Although Ebs1 has been shown to be involved in NMD it is not a core factor which may explain why deletion of *TMA20* or *TMA22* in a *cdc13-1 ebs1Δ* background increases fitness, in contrast to deletion of *TMA20* or *TMA22* in *cdc13-1 nmd2Δ* (Luke et al., 2007).

Deletion of NMD genes suppresses the *cdc13-1* temperature sensitive phenotype, in part, by up-regulating Stn1 and Ten1 levels (Holstein et al., 2014). One possible explanation, therefore, to explain how *TMA20* and *TMA22* act in the same pathway as NMD to affect fitness of *cdc13-1* is that *tma20Δ* and *tma22Δ* also increase levels of Stn1 and Ten1. It would be of interest to determine if Tma20/Tma22 affect levels of Stn1 and Ten1 and if they do, the mechanism by which they inhibit Stn1 and Ten1. It would also be interesting to further explore the relationship between Tma20/Tma22

and NMD to test if their biochemical functions overlapped. NMD has been shown to promote ribosome recycling on premature termination codons so it is possible that both Tma20/Tma22 affect levels of the same transcripts by promoting ribosome splitting following translation termination (Ghosh et al., 2010). Additionally deletion of *TMA20* and NMD factors both result in an increase in stop codon read-through suggesting that increase in stop codon read-through may promote fitness of *cdc13-1* (Fleischer et al., 2006a).

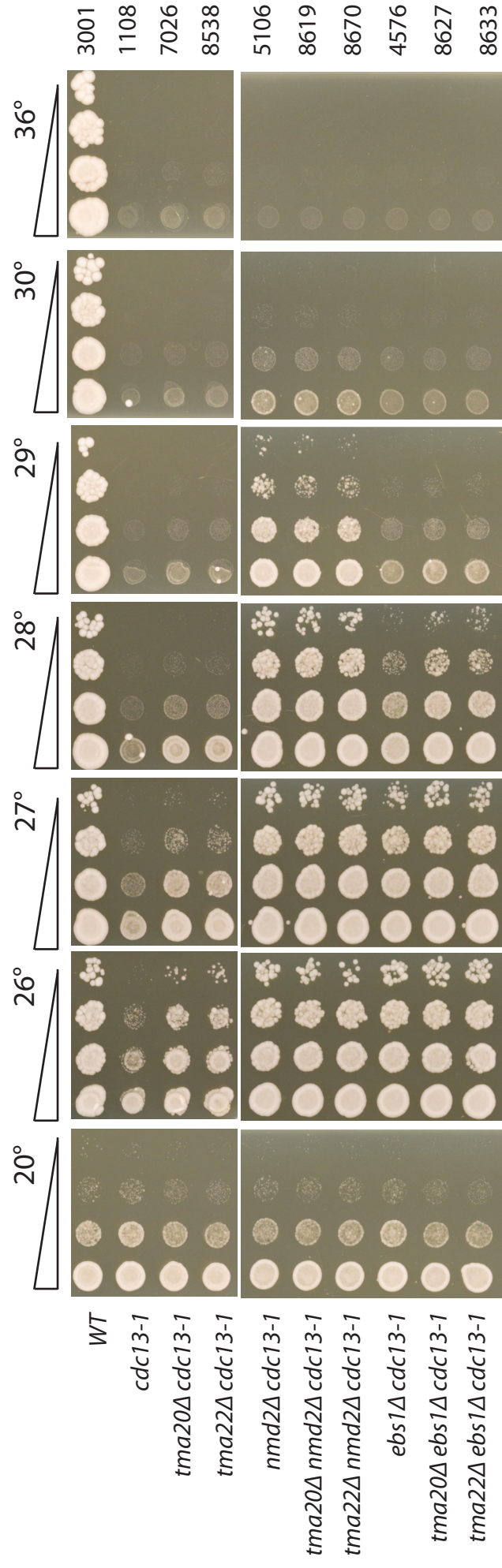


Figure 15: *TMA20* and *TMA22* act in the same pathway as *NMD* to affect fitness of *cdc13-1*. Yeast strains of the indicated genotypes were grown to saturation in liquid YEPD, serially diluted, 5-fold, in water, and spotted onto YPD solid media. Plates were incubated for two days at the indicated temperatures before being photographed. Genotypes and strain numbers are listed on either side of the spot tests.

### 3.6 Evidence that increase in levels of stop codon read-through does not suppress *cdc13-1*

It has been shown that *tma20Δ* has a 20-fold increase in the level of stop codon read-through compared with wild type strains (Fleischer et al., 2006a). To examine if an increase in the level of stop codon read-through might contribute to the loss of fitness of *cdc13-1* we took advantage of the PSI<sup>+</sup> prion status of W303 strains. Prions are infectious isoforms of normal proteins, which exist in yeast and mammals and cause diseases such as Bovine spongiform encephalopathy (BSE) (Lloyd et al., 2013). eRF3, encoded by *SUP35*, can exist in a prion state known as PSI<sup>+</sup> which compromises the normal function of eRF3 as a termination factor (Serio and Lindquist, 1999). Therefore PSI<sup>+</sup> have increased stop codon read-through compared with PSI<sup>-</sup> (Keeling et al., 2004). PSI<sup>+</sup> can be converted to PSI<sup>-</sup> by growing cells in low level of guanidine hydrochloride (Eaglestone et al., 2000) and the PSI<sup>+</sup> status of strains can be determined using a GFP labelled version of eRF3 expressed from a plasmid (Greene et al., 2009). Since the amino (N) and highly charged middle (M) domains of eRF3 were fused with GFP to create the labelled version of eRF3 the fusion protein is referred to as *SUP35NM* (Greene et al., 2009). In PSI<sup>+</sup> cells transformed with the *SUP35NM* reporter plasmid, discrete GFP foci can be observed due to the eRF3-GFP fusion protein binding eRF3 prion aggregates (Greene et al., 2009).

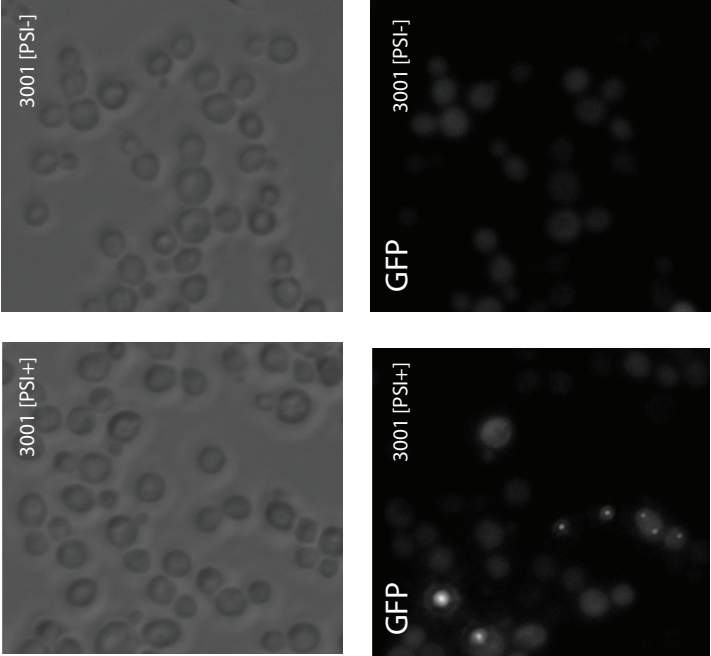
To convert PSI<sup>+</sup> strains to PSI<sup>-</sup> WT strains were grown on YPD solid media containing 3mM GuHCl at 23°C for 3 days. To confirm PSI<sup>+</sup> strains were converted to PSI<sup>-</sup> WT, *cdc13-1*, *cdc13-1 tma22Δ* and *cdc13-1 tma20Δ rad9Δ* grown in the presence or absence of GuHCl were transformed with *SUP35NM* reporter plasmid and 2- 4 of the resulting colonies were checked for the presence of GFP foci. There were no GFP foci detected in any of the strains grown on GuHCl whereas there was about 1 focus per 100 cells detected in cells not exposed to GuHCl (Figure 16a).

To check if *cdc13-1* lethality was affected by the PSI status fitness of *cdc13-1* [PSI<sup>+</sup>] and *cdc13-1* [PSI<sup>-</sup>] were compared by spot test. There was no difference in the growth of *cdc13-1* [PSI<sup>+</sup>] compared with *cdc13-1* [PSI<sup>-</sup>] at any of the temperatures tested suggesting that the PSI status of cells does not affect fitness of *cdc13-1* (Figure 16b). Similarly there was no difference in the fitness of *cdc13-1 tma22Δ* [PSI<sup>+</sup>] and *cdc13-1 tma20Δ rad9Δ* [PSI<sup>+</sup>] compared with *cdc13-1 tma22Δ* [PSI<sup>-</sup>] and

*cdc13-1 tma20Δ rad9Δ* [PSI<sup>-</sup>] respectively, providing further evidence that the PSI status of W303 does not affect fitness of *cdc13-1* (Figure 16b).

The finding that PSI status does not affect fitness of *cdc13-1* strongly suggests that inefficient translation termination, which leads to increased stop codon read-through, does affect fitness of *cdc13-1*. Further *tma20Δ*, *hcr1Δ* and *nam7Δ*, each of which has been shown to result in increased stop codon read-through phenotype, are unlikely to promote fitness of *cdc13-1* by increasing levels of stop codon read-through (Fleischer et al., 2006a, Beznoskova et al., 2013).

A



73  
B

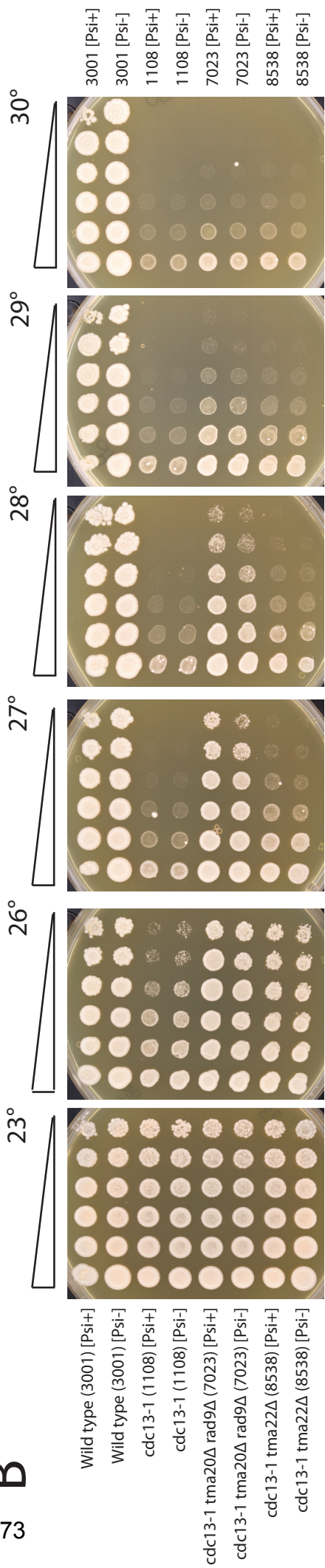


Figure 16: Evidence that increase in levels of stop codon read-through does not suppress *cdc13-1*. A) PSI status of strains was determined before and after growth on 3mM guanidine hydrochloride by detection of fluorescent aggregates. B) Yeast strains of the indicated genotypes were grown to saturation in liquid YEPD, serially diluted, 5-fold, in water, and spotted onto YPD solid media. Plates were incubated for two days at the indicated temperatures before being photographed. Genotypes and strain numbers are listed on either side of the spot tests.



### 3.7 *TMA20* and *TMA22* act in parallel pathways to DNA damage genes

In mammalian cells *TMA20* (MCT-1) overexpression or knockdown has been shown to affect the level of many DNA damage genes (Nandi et al., 2007). Many of the gene deletions which allow *cdc13-1* strains to evade cell cycle arrest are those involved in DNA damage response, since they evade cell cycle checkpoint (Addinall et al., 2008). We therefore hypothesised that *TMA20* and *TMA22* may decrease fitness of *cdc13-1* by affecting levels of some DNA damage response proteins. Rad9, Exo1 and Rad24 are all essential for cell cycle arrest following telomere uncapping whereas cell cycle arrest is partially dependant on Chk1 (depicted in Figure 3)(Gardner et al., 1999). Exo1 is a nuclease, which activates cell cycle arrest by generating ssDNA in *cdc13-1* (Zubko et al., 2004), whereas Rad24, Rad9 and Chk1 are checkpoint protein. Rad24 stimulates cell cycle arrest by loading of the checkpoint complex 9-1-1 (Rad17, Mec3, and Ddc1) onto damaged telomeres. Rad9 is important for activation of effector kinases Rad53 and Chk1, which ultimately leads to G2/M arrest (Vialard et al., 1998a). Phosphorylation targets of Chk1, include Pds1 (Securin) and DNA damage signal transducer Mec1.

To test the hypothesis that *Tma20* and *Tma22* affect the fitness of *cdc13-1* by regulating levels of DNA damage response proteins *tma20Δ* and *tma22Δ* were crossed with *cdc13-1 rad9Δ rad24Δ exo1Δ* strain, diploids were sporulated and all possible combinations of haploids were created. In addition *cdc13-1 tma20Δ* and *cdc13-1 tma22Δ* were combined with *chk1Δ*. *TMA20* and *TMA22* were found to act in a parallel pathway to *CHK1* since, in a *cdc13-1* background, *tma20Δ chk1Δ* and *tma22Δ chk1Δ* grew better than any of the respective single deletions at 27°C, 28°C and 29°C (Figure 17). A similar effect is observed when *cdc13-1 tma20Δ* and *cdc13-1 tma22Δ* are combined with *rad9Δ*, *exo1Δ* and *rad24Δ* (Figure 17). In a *cdc13-1* background *tma20Δ* and *tma22Δ* enhance the growth of *rad9Δ* at 28°C and 29°C, *exo1Δ* 28°C and 29°C and *rad24Δ* at 29°C and 30°C, suggesting that *TMA20* and *TMA22* act in parallel pathways to *RAD9*, *EXO1* and *RAD24* to decrease fitness of *cdc13-1* (Figure 17). In agreement with a model where *TMA20* and *TMA22* act in a parallel pathway to *RAD24* and *RAD9* and *EXO1* to suppress *cdc13-1*, *tma20Δ* and *tma22Δ* improves growth of *cdc13-1 rad24Δ rad9Δ* at 28°C, 29°C and 30°C, *cdc13-1 rad24Δ exo1Δ* at 29°C and 30°C, *cdc13-1 rad9Δ exo1Δ* at 30°C and *cdc13-1 rad9Δ rad24Δ exo1Δ* at 29°C and 30°C.

The finding that Tma20 and Tma22 act in parallel pathways to *RAD9*, *RAD24*, *EXO1* and *CHK1* to decrease fitness of *cdc13-1* implies that Tma20 and Tma22 do not contribute to the loss fitness of *cdc13-1* by influencing levels of DNA damage response proteins. It also indicates that in yeast *TMA20* may not function in the DNA damage response as MCT-1 does (Hsu et al., 2007). This is consistent with our hypothesis that Tma20 and Tma22 affect *cdc13-1* by modulating levels of other telomere capping proteins, such as Stn1 and Ten1.

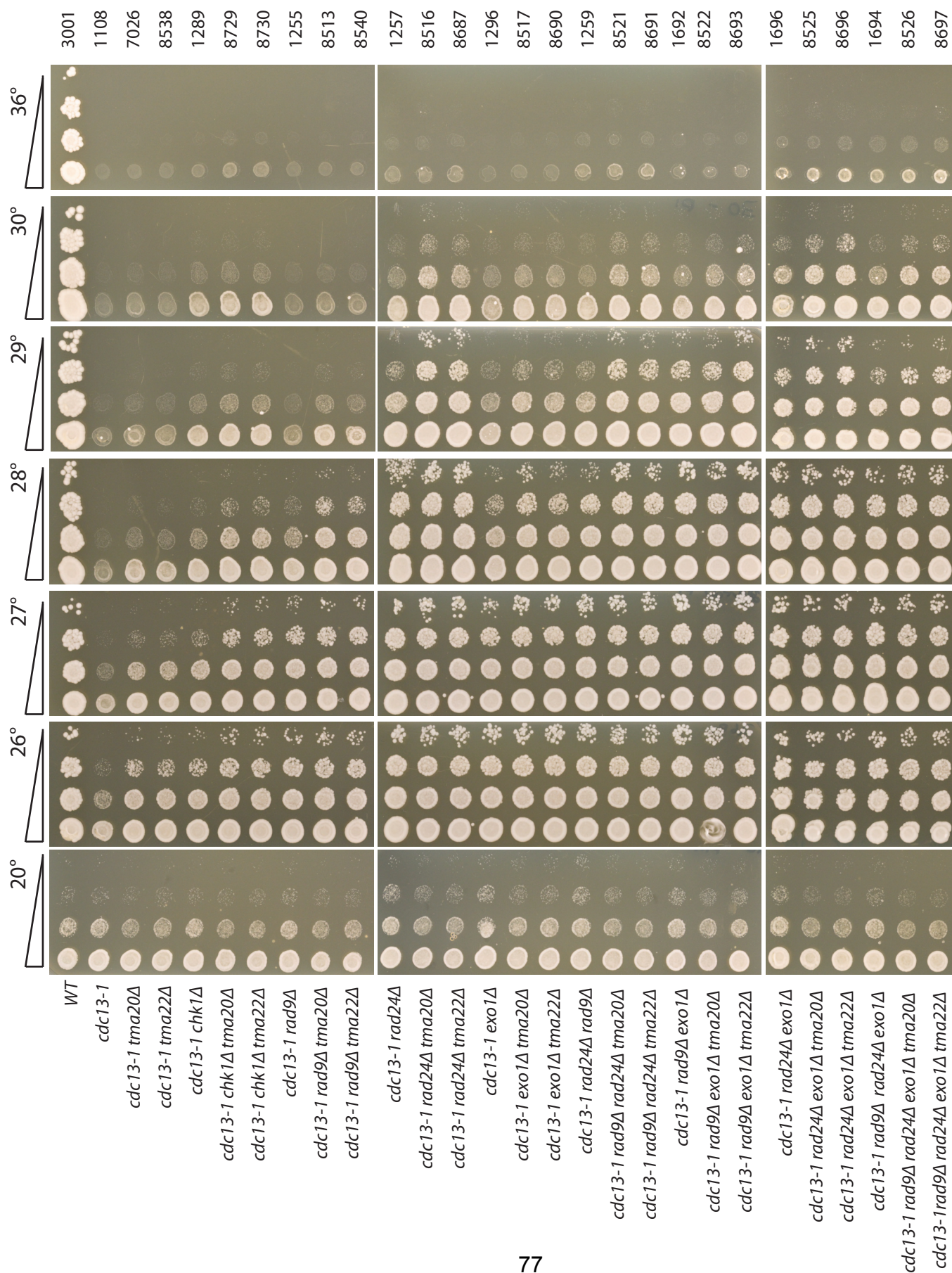


Figure 17: *TMA20* and *TMA22* act in parallel pathways to DNA damage genes. Yeast strains of the indicated genotypes were grown to saturation in liquid YEPD, serially diluted, 10-fold, in water, and spotted onto YPD solid media. Plates were incubated for two days at the indicated temperatures before being photographed. Genotypes and strain numbers and listed on either side of the spot tests.

### 3.8 Mutations in conserved phosphorylation sites affect function of the Tma20

The level of MCT-1 increases in response to  $\gamma$ -irradiation, implying that it functions to mediate the cellular response to DNA damage (Herbert et al., 2001). The induction of MCT-1 upon DNA damage was subsequently shown to be dependent on phosphorylation by p44/p42 MAPK (ERK) (Nandi et al., 2007). We hypothesised that phosphorylation sites are conserved and since Tma20 decreases fitness of *cdc13-1* we predict that levels of Tma20 might increase as a response to telomeres becoming damaged.

To examine the conservation of amino acids between yeast, Drosophila and humans *TMA20* was aligned to its' human and Drosophila homologs, MCT-1 (Figure 18a). The phosphorylation site required for cell growth in mammalian cells (T81) and a phosphorylation site important for proper function of the Drosophila homolog of MCT-1 (T118/S119) are both present in yeast *TMA20* suggesting that either could perform an important role in yeast (Figure 18a).

To explore the importance of phosphorylation sites to the function of Tma20, two new alleles of *TMA20* were created where phosphorylation sites were abolished by point mutation. *TMA20 T118A S119A* contains two point mutations in the proposed Cdk1 (*CDC28*) phosphorylation site whereas *TMA20 S81A* harbours a point mutation in the proposed MAPK phosphorylation site. Tma20 was expressed from a centromeric plasmid, transformed into *TMA20* null strains. Since *tma20 $\Delta$*  increases fitness of *cdc13-1* we reasoned that an allele of *TMA20* resulting in loss of Tma20 function would also increase fitness of *cdc13-1*.

While *cdc13-1 tma20 $\Delta$*  grew better than *cdc13-1* at 26°C and 27°C, growth of *cdc13-1 tma20 $\Delta$*  cells transformed with the plasmid expressing *TMA20* was indistinguishable from growth of *cdc13-1* demonstrating that expression of *TMA20* from a plasmid is sufficient rescue the deletion phenotype of *tma20 $\Delta$*  in *cdc13-1* and also confirms that the suppression of *cdc13-1* by *tma20 $\Delta$*  is not a result of a neighbouring gene effect (Figure 18b and c).

*cdc13-1 tma20 $\Delta$*  transformed with the plasmid expressing *TMA20 S79A* grew better than *cdc13-1 tma20 $\Delta$*  transformed with the plasmid expressing *TMA20* but not as well as *cdc13-1 tma20 $\Delta$*  transformed with vector only suggesting that S79A partially

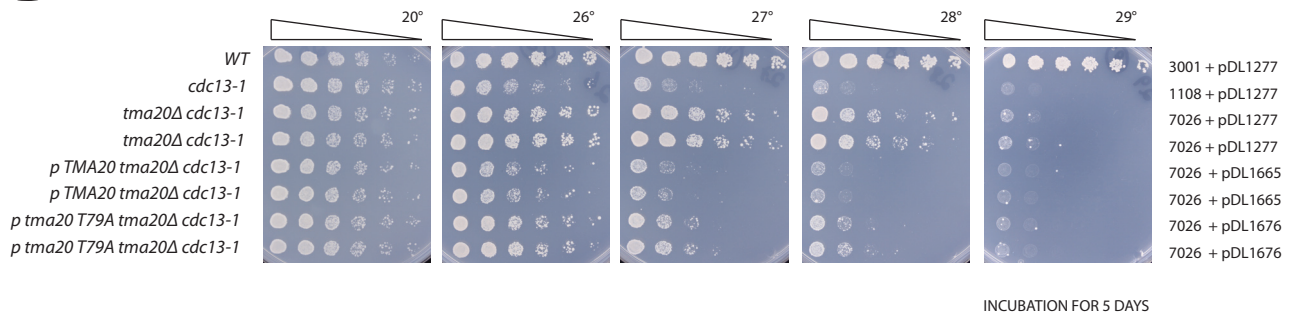
disrupts the activity of *TMA20* (Figure 18b). *cdc13-1 tma20Δ* transformed with the plasmid expressing *TMA20 T118A S119A* also grew better, at 26°C, than *cdc13-1 tma20Δ* transformed with the plasmid expressing *TMA20* but not as well as *cdc13-1 tma20Δ* transformed with vector only suggesting that T118 S119 phosphorylation sites are also required for proper function of *TMA20* but *TMA20* retains some of its activity without this phosphorylation site (Figure 18c).

The finding that Tma20 phosphorylation sites are important for its proper function is consistent with our hypothesis that Tma20 is phosphorylated and stabilised in response to DNA damage. However, it is also possible that mutations introduced into the proposed phosphorylation sites disrupt Tma20 function simply by preventing proper folding of Tma20 or disrupting Tma20 interactions with other proteins or RNA. It would be of interest therefore to observe phosphorylation or change in the level of Tma20 following DNA damage.

# A

Tma20 (Yeast)	MFKKFTRED-VHSRSKVKSSIQRTLKAKLVKQYPKIEDVIDELIPKKSQIELIKCEDKIQ
MCT-1 (Human)	MFKKFDEKENVSNICQLKTSVIKGIKNQLIEQFPGIEPWLNQIMPKKDPVKIVRCHEHIE
MCT-1 (Drosophila)	MFKKFEEKDSISSIQQLKSSVQKIRAKLLEAYPKLESHIDLILPKKDSYRIAKCHDHIE
	***** .::: . . :*:*: : : :*:*: :* :* : : :*:*. . :*:*:*:
	MAPK phosphorylation site CDK1 phosphorylation site
Tma20 (Yeast)	LYS-VDGEVLFFQK-FDELIP <sup>S</sup> LKLVHKFPEAYPTVQVDRGAIKFVLSGANIMCPGL <sup>TS</sup> A
MCT-1 (Human)	ILT-VNGELLFFRQREGPFYP <sup>T</sup> LRLLHKYPFILPHQQVDKGAIKFVLSGANIMCPGL <sup>TS</sup> P
MCT-1 (Drosophila)	LLLNGAGDQVFFRHRDGPWMP <sup>T</sup> LRLLHKFPYFVTMQQVDKGAIRFVLSGANVMCPGL <sup>TS</sup> P
	: * : :*:*: :*:*:*:*:* ***:***:*****:*****
Tma20 (Yeast)	GADLPPAPGYEKGITIVVINAENKENALAI GELMMGTEEEKSVNKGHSIELIHHLGDPLWN
MCT-1 (Human)	GAKLYPAA---VDTIVAIMAEGKQHALCVGMKMSAEDIEKVNKGIGIENIHYLNDGLWH
MCT-1 (Drosophila)	GACMTPAD---KDTVVAIMAEGKEHALAVGLLTLSTQEI LAKNKGIGIETYHFLNDGLWK
	** : ** *:*. * * *:*.*:*: : : :*:*: * * * . * * * . * * *
Tma20 (Yeast)	FSVE-
MCT-1 (Human)	MKTYK
MCT-1 (Drosophila)	SKPVK

# B



# C

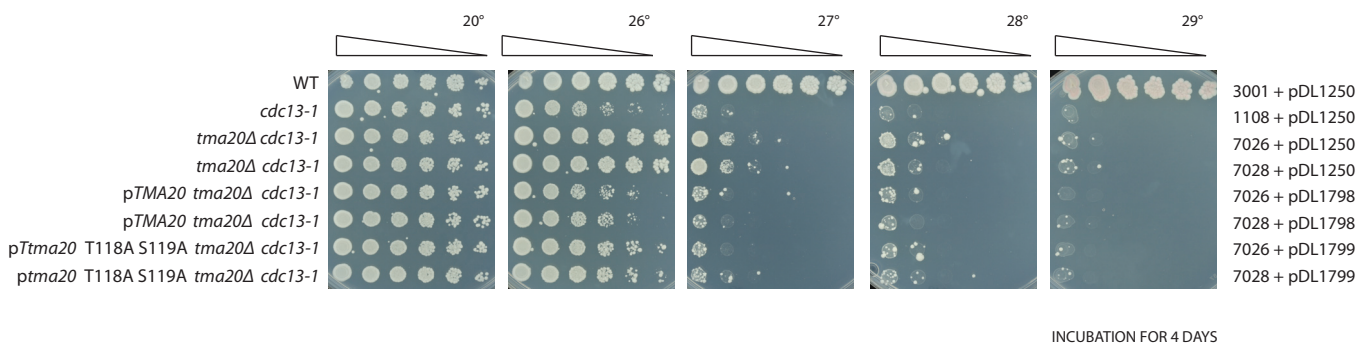


Figure 18: Mutations in conserved phosphorylation sites affect function of the Tma20.

A) Sequence alignment of *TMA20* from *Saccharomyces cerevisiae* with MCT-1 from human and *Drosophila*. B and C) *TMA20*, *tma20 T79A* and *tma20 T118A T119A* were expressed from a centromeric plasmid which was transformed into the indicated yeast strains. Cells were grown to saturation in liquid SD media, serially diluted, 5-fold, in water, and spotted onto solid SD media. Plates were incubated for 5 (B) and 3 (C) days at the indicated temperatures before being photographed



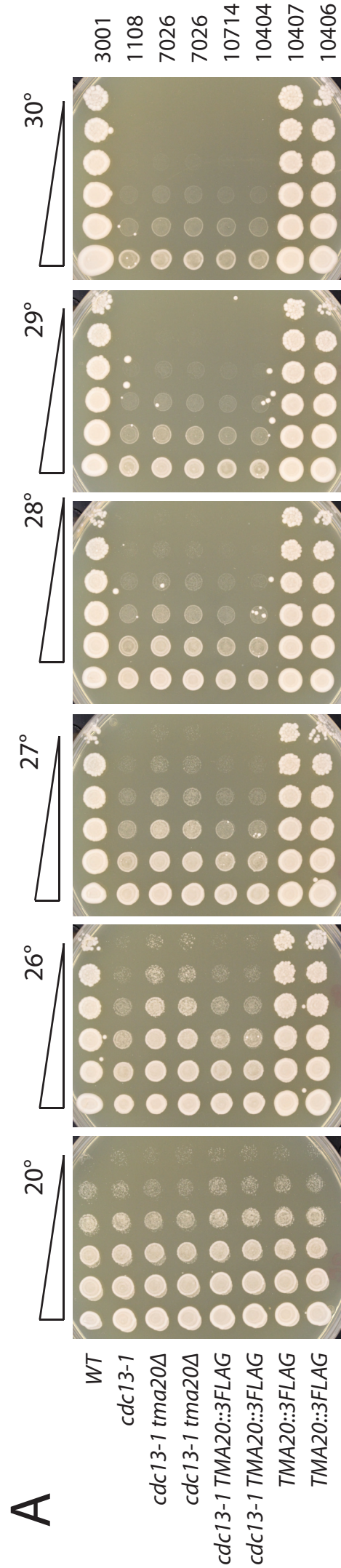
### 3.9 No evidence that Tma20 is up regulated or phosphorylated following telomere uncapping

Since we found that both conserved phosphorylation sites (*T81A* and *T118A S119A*) are important for the proper function of Tma20 we decided to test if we could detect a change in the phosphorylation status of Tma20 following telomere uncapping. We hypothesised that Tma20 is phosphorylated following telomere uncapping, which would lead to an increase in protein level, due to it being stabilised. To test this we sought to measure levels of Tma20 in cells before and after *cdc13-1* induced telomere damage. If Tma20 is phosphorylated this may be visible as a change in the molecular weight or the appearance of a second band running slightly above Tma20.

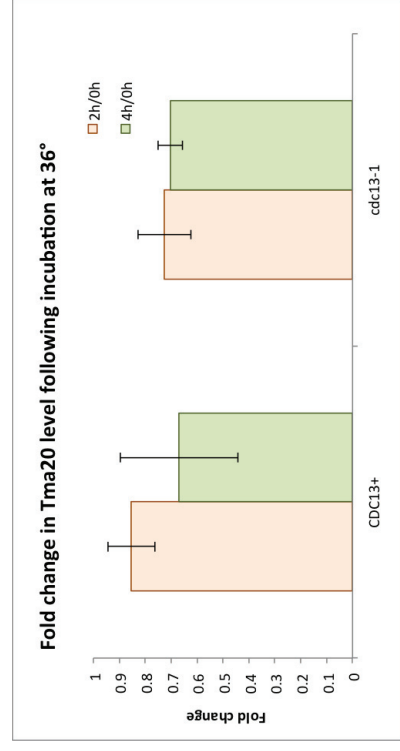
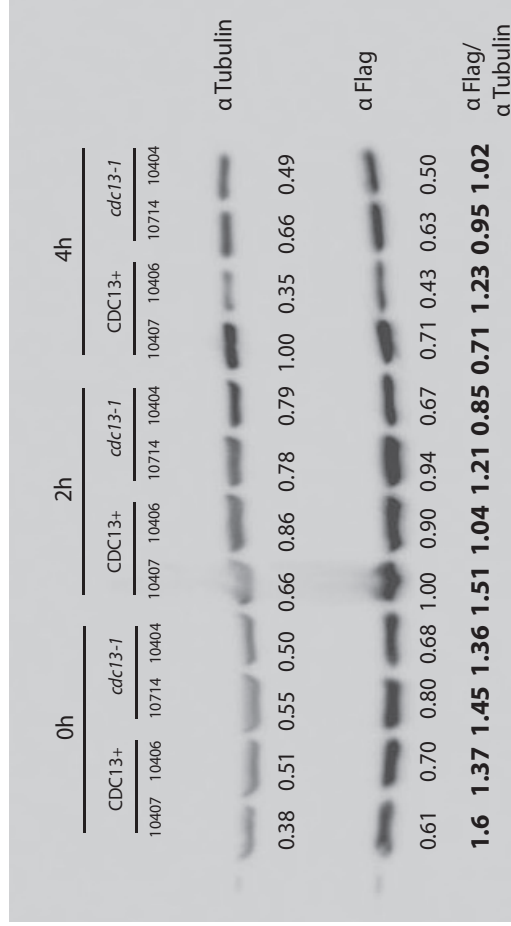
To measure the abundance of Tma20 a FLAG epitope tag was fused to the C-terminus of Tma20. To verify that the FLAG epitope tag does not affect the function of Tma20 we compared growth of *cdc13-1 tma20Δ* with *cdc13-1 TMA20-3FLAG*. While at 26°C-28°C *cdc13-1 tma20Δ* grew better than *cdc13-1*, growth of *cdc13-1 TMA20-3FLAG* was indistinguishable from growth of *cdc13-1* indicating that the function of Tma20 is intact when fused with FLAG (Figure 19a).

To test if the level of Tma20 increased following exposure to telomere damage, levels of Tma20 were measured in *cdc13-1* and WT cells grown to exponential phase at 23°C and then shifted to 36°C for 0, 2 or 4 hours. There was a general trend to a decrease in the level of Tma20 when cells were shifted to 36°C however because this was observed in both *cdc13-1* and WT cells it was not attributed to telomere uncapping (Figure 19b). There was also no change in the molecular weight of Tma20 detected, which would indicate a post-translational modification such as phosphorylation (Figure 19b). A phosphate-affinity polyacrylamide gel electrophoresis would be required to confirm that Tma20 is not phosphorylated in response to telomere uncapping.

Our data indicate Tma20 does not play an active role in responding to DNA damage, induced by *cdc13-1*. It is unclear however if levels of Tma20 increase following other types of DNA damage.



**B**



**Figure 19 - No evidence that Tma20 is upregulated or phosphorylated following telomere uncapping**

Figure 19: No evidence that Tma20 is upregulated or phosphorylated following telomere uncapping. A) Yeast strains of the indicated genotypes were grown to saturation in liquid YEPD, serially diluted, 5-fold, in water, and spotted onto YPD solid media. Plates were incubated for two days at the indicated temperatures before being photographed. Genotypes and strain numbers are listed on either side of the spot tests. B) Cells were cultured at 23°C until mid log phase (0.6OD) was achieved and then incubated at 36°C for indicated time. Whole cell extracts were ran on a gel, which was probed using anti-Flag to detect Tma20 and anti-tubulin as a loading control (C) Quantification of B using Image J to quantify band intensities.

### 3.10 Discussion

In this chapter we confirm that *TMA20* and *TMA22* null alleles increase fitness of cells with mutations in CST complex members, *cdc13-1* and *stn1-13*. We also showed that defects in recycling of post termination ribosomes induced by *hcr1Δ* also decreases the fitness of *cdc13-1*, in a manner that is rescued by overexpression of the ribosome-recycling factor, Rli1. Our results indicate that *TMA20* and *TMA22* function in the same pathway as NMD, but different pathways to *HCR1* and DNA damage response genes *RAD24*, *RAD9*, *EXO1* or *CHK1* to increase the fitness of *cdc13-1*. We also show that conserved phosphorylation sites on Tma20 are required for Tma20 to decrease fitness of *cdc13-1* but found no evidence that Tma20 is phosphorylated or up-regulated in response to telomere uncapping.

#### 3.10.1 Interacting partners, Tma20 and Tma22 have a novel role in telomere biology

We confirm high throughput QFA data showing that *TMA20* and *TMA22* null alleles increase fitness of cells with *cdc13-1* (Addinall et al., 2011) and *stn1-13* (Holstein et al., 2017) induced damage (Figure 11a and b). In every experiment that we conducted *TMA20* deletion phenotypes were indistinguishable from *tma20Δ* and *tma20Δ tma22Δ* in agreement with the idea Tma20 and Tma22 function as a complex (Figure 11, Figure 13, Figure 15 and Figure 17) (Fleischer et al., 2006a). This idea is further supported by recent structural analysis showing that mammalian homologs of Tma20 and Tma22 function as a heterodimer (Lomakin et al., 2017). We also confirmed QFA data showing that, in contrast with some genes such as checkpoint genes *RAD9* and nucleases *EXO1*, deletion of *TMA20* does not also reduce fitness of cells with *ku70Δ* telomere defects (Figure 11c) (Addinall et al., 2011). This indicates that Tma20 and Tma22 may have a specific rather than general role in telomere biology, for example by regulating levels of a specific capping protein rather inducing a checkpoint response. In support of the idea that Tma20 and Tma20 do not have a major role in telomere biology, the telomeres of *tma20Δ* are indistinguishable from wild type (Figure 12).

#### 3.10.2 Decreased recycling of post termination ribosomes improves *cdc13-1* fitness

Interestingly we showed that *hcr1Δ* is stronger suppressor of *cdc13-1* thermo sensitivity than deletion of *tma20Δ* or *tma22Δ* (Figure 13). The homologs of both

Tma20/Tma22 and Hcr1 promote release of ribosome subunits following translation termination (Figure 9) (Skabkin et al., 2010a, Pisarev et al., 2007). Surprisingly however we discovered that Tma20 and Hcr1 function in different pathways to decrease fitness of *cdc13-1* suggesting that their functions in translation are distinct (Figure 13). Consistent with this we showed that overexpression of recycling factor Rli1, in the context of *cdc13-1*, suppresses *hcr1Δ* but not *tma20Δ* (Figure 14). This suggests the increase in fitness of *cdc13-1* observed upon deletion of *HCR1* is explained by a decrease in ribosome recycling but that Tma20 decreases fitness of *cdc13-1* by a different mechanism. Alternatively overexpression of Rli1, which promotes 60S dissociation from the mRNA following translation termination (Figure 9) (Becker et al., 2012) (Shoemaker and Green, 2011), may compensate for the specific ribosome-recycling defect of *hcr1Δ* but not *tma20Δ*. In agreement with this, Hcr1 is implicated in promoting 60S dissociation from the mRNA in addition to the release of de-acetylated tRNA and 40s from mRNA (Beznoskova et al., 2013), whereas homologs of Tma20/Tma22 have only been shown to promote the release of de-acetylated tRNA and 40S from mRNA (Figure 9) (Skabkin et al., 2010a). It is unclear if Tma20 decreases fitness of *cdc13-1* by promoting the release of de-acetylated tRNA and 40s from mRNA or via a different function.

Hcr1, Tma20/Tma22 and Nmd2, all associate with termination machinery and have been shown to promote the release of ribosomal subunits from the transcript (Beznoskova et al., 2013) (Skabkin et al., 2010a, Ghosh et al., 2010). The observation therefore that *hcr1Δ*, *tma20Δ* and *nmd2Δ* each improve fitness of *cdc13-1* highlights the importance of ribosome recycling as a mechanism of regulating gene expression. We speculate that some mRNAs will be particularly susceptible to decreases in ribosome recycling efficiency. Transcripts that contain uORFs may be among the mRNAs that decrease in abundance when ribosomes are inefficiently recycled. This is because ribosomes that do not dissociate from the TL after translation of a uORF will block the path of other scanning 40S complexes and can result in the transcript being degraded by NMD (Figure 7d) (Gaba et al., 2005a).

### 3.10.3 *Tma20 and Tma22 function in the same pathway as NMD to decrease fitness of cdc13-1*

We discovered that *TMA20* and *TMA22* function in the same pathway as *NMD2* and parallel pathways to DNA damage response genes *RAD24*, *RAD9*, *EXO1* and *CHK1*

to decrease fitness of *cdc13-1* (Figure 15 & Figure 17). NMD is thought to decrease the fitness of *cdc13-1*, at least in part, by decreasing the levels of other telomere capping proteins, Stn1 and Ten1 (Addinall et al., 2011). *RAD24*, *RAD9*, *EXO1* and *CHK1* on the other hand decrease the fitness of *cdc13-1* since they contribute to the DNA damage response pathways that cause cell cycle arrest in *cdc13-1* (Figure 3) (Lydall, 2009). This led us to hypothesis that Tma20 and Tma22 modulate levels of telomere capping proteins such as Stn1 and Ten1, rather than DNA damage response or checkpoint proteins to affect fitness of *cdc13-1*.

Some of our data however, is in conflict with the idea that *NMD* and *TMA20* act in the same pathway to affect general telomere fitness. *nmd2Δ* is an enhancer of *yku70Δ* temperature sensitivity whereas deletion of *TMA20* has no effect on growth of *ku70Δ* at any temperature (Addinall et al., 2011). Additionally *nmd2Δ* have short telomeres (Dahlseid et al., 2003) whereas *tma20Δ* telomeres are normal (Figure 12). One explanation for these observations could be that the roles of *TMA20* and *NMD2* at telomeres only partially overlap. Alternatively *TMA20* may only slightly increase the levels of telomere capping proteins such as Ten1 and Stn1 whereas *nmd2Δ* results in a large increase in the level of Ten1 and Stn1.

#### *3.10.4 Increased stop codon read-through reported in tma20Δ, nmd2Δ and hcr1Δ does not explain the increase in cdc13-1 fitness observed in these strains*

Interestingly *tma20Δ*, *nmd2Δ* and *hcr1Δ* have each been shown to increase stop codon read-through (Fleischer et al., 2006a, Beznoskova et al., 2013) (Keeling et al., 2004, Altamura et al., 2016), which made us speculate that they may improve fitness of *cdc13-1* by increasing stop codon read-through. Several heritable cancers have been associated with mutations that create premature termination codons and therapies that promote stop codon read-through have been developed to restore normal expression of PTC containing transcripts (Bordeira-Carrico et al., 2012). It is of relevance therefore to gain more knowledge about both the proteins involved in stop codon read-through and the consequences of defective stop codon read-through (Bordeira-Carrico et al., 2012). W303 yeast strains used in this study contain a prion form of eRF3, which interferes with normal translation termination and is referred to as [PSI<sup>+</sup>] (Serio and Lindquist, 1999). We took advantage of this, and the fact that [PSI<sup>+</sup>] strains can be converted to [PSI<sup>-</sup>] but found no difference in the fitness

of *cdc13-1* strains with higher [PSI<sup>+</sup>] or lower [PSI<sup>-</sup>] levels of stop codon read-through (Figure 16).

#### *3.10.5 No evidence to suggest that Tma20 actively responds to cdc13-1 induced DNA damage*

We explored the possibility that Tma20 might actively respond to telomere damage induced by *cdc13-1*. In response to irradiation treatment the mammalian homolog of Tma20 has been shown to become more stable by phosphorylation of conserved sites (Herbert et al., 2001). In contrast with this we did not observe any phosphorylation or increase in abundance of Tma20 upon telomere uncapping (Figure 19). This is surprising since irradiation induces DSBs and the DNA damage response to a DSB is similar to that of uncapped telomeres, (Lydall, 2009) highlighting the different functions of Tma20 and MCT-1. However it would be necessary to conduct further experiments to confirm that Tma20 is not phosphorylated such as running a phosphorylation gel that would be more sensitive to detecting phosphorylation.

Interestingly however we did find that the conserved phosphorylation sites of Tma20 are partially required for Tma20 to decrease the fitness of *cdc13-1* (Figure 18). It is unclear if this is a result of these particular amino acids being important for proper protein folding. We could test this by treating cells with phosphorylation inhibitors and measuring the ability of Tma20 to decrease fitness of *cdc13-1*.

## **Chapter 4. No Evidence that Tma20 and Tma22 promote translation re-initiation as in *Drosophila* and mammalian cells**

### **4.1 Identification of genes with uORFs**

In *Drosophila* the homologues of Tma20 (MCT-1) and Tma22 (DENR) are required for optimal translation of genes that contain upstream open reading frames (uORFs) by promoting translation re-initiation (Schleich et al., 2014b). In this study the dependence of a gene on MCT-1 and DENR for optimal translation correlated with the strength of the uORF initiation context (the Kozak score) and inversely correlated with the uORF length (Schleich et al., 2014b). A more recent publication showed that mammalian homologs of Tma20 and Tma22 are also required for the optimal translation of genes that contain uORFs in strong initiation contexts. In contrast with *Drosophila* however, the mammalian homologs of Tma20 and Tma22 only promote translation of genes that have uORFs of a single amino acid in length (Schleich et al., 2017).

To test the hypothesis that Tma20 and Tma22 promote translation of yeast genes that contain uORFs a bioinformatics analysis of yeast TL sequences was conducted using recently available yeast TL length predictions (Nagalakshmi et al., 2008) (Arribere and Gilbert, 2013). All uORFs, including those with a stop codon located in the main ORF, termed overlapping ORFs, were identified. To predict the likelihood that a uORF initiation codon would be recognised by a 43S ribosome scanning the TL, the similarity of the sequence flanking initiation codons and the consensus Kozak sequence was determined and expressed as a 'Kozak score'. This was accomplished using a position weight matrix of the consensus Kozak sequence, created by counting the occurrences of each nucleotide at positions -4, -3, -2, -1, +1 and +2 relative to the ATG of the main translation start site of each gene in the yeast genome (Figure 20a). The Kozak score of a given uORF was then generated as in (Schleich et al., 2014b), using a multiplicative model whereby the frequency of occurrences of each nucleotide (obtained from the Kozak position weight matrix) from position minus 4 to plus 2 relative to the ATG were multiplied together. For example the consensus Kozak sequence of AAAAatgAC would yield a Kozak Score of  $[0.42902 \times 0.576061 \times 0.397639 \times 0.439928 \times 0.317693 \times 0.366557]$  which is 0.005035. A high 'Kozak score' reflects an initiation context similar to the Kozak consensus



sequence. The number of uORFs per gene was also counted and a 'total Kozak score' calculated as the sum of all the Kozak scores from each of the uORFs.

The sequences of 6600 genes were retrieved from SGD. TL sequences of 5164 genes were obtained from TL length estimates previously published (Nagalakshmi et al., 2008) (Arribere and Gilbert, 2013). From this the locations of all uORF and oORFs were determined and their 'Kozak scores' calculated. In total 1663 and 464 uORF and oORFs were identified respectively (Figure 20b). 565 genes contain one or more uORFs and 374 genes contain one or more oORFs (Figure 20b). The proportion of uORFs identified (12.75%) is in agreement with previous estimates of 13% (Lawless et al., 2009). oORFs were found to be less common than uORFs, only occurring in 7.36% of genes. Data from the bioinformatics analysis is summarised in Figure 20b and a screenshot of the results are displayed in Figure 20c.

A



	-4	-3	-2	-1	0	0	0	1	2
A	0.42902	0.576061	0.397639	0.439928	1	0	0	0.317693	0.265093
C	0.197699	0.094441	0.21832	0.178123	0	0	0	0.132247	0.366557
G	0.148984	0.189331	0.143903	0.164824	0	0	1	0.288255	0.151973
T	0.224298	0.140167	0.240137	0.217125	0	1	0	0.261805	0.216378

B

SystematicName	GeneName	uAUG Kozak score	frame	length of uORF	TL_Length	ORF Kozak score	length of oORF
YAL008W	FUN14	0	0	0	49	24.86245063	0
YBR255W	MTC4	2.281615149	OUT	0	24	23.20025119	24
YGR164W		0	0	0	unknown	11.13464007	0
YGR131W	FHN1	0.515898558	IN	12	180	7.471118567	0
YGR131W	FHN1	0.192696189	OUT	0	180	7.471118567	111
YGR131W	FHN1	1.57669464	IN	15	180	7.471118567	0
YPL144W	POC4	0	0	0	27	0.369833358	0
YBR135W	CKS1	0	0	0	105	1.270021089	0
YBR160W	CDC28	0	0	0	79	3.603764904	0
YJL082W	IML2	0	0	0	119	3.186460693	0
YJL142C	IRC9	0	0	0	unknown	23.20025119	0
YPL191C		8.485595644	OUT	42	101	1.104699674	0
YGL215W	CLG1	0	0	0	464	16.53163589	0
YKL074C	MUD2	0	0	0	29	2.150882908	0
YLR467C-A		0	0	0	unknown	2.411318319	0
YJL077C	ICS3	7.926617413	OUT	15	28	1.493194493	0
YKL096W-A	CWP2	0	0	0	49	15.15656124	0
YIL124W	AYR1	0	0	0	32	2.671013028	0

C

	Number
Total Genes	6600
genes with defined TL lengths	5146
uAUGs found	2127
uORFs found	1663
oORFs found	464
uORFs in frame	1088
uORFs out of frame	575
oORFs in frame	46
oORFs out of frame	418
genes with one or more uORF	656
genes with one or more oORF	379

Figure 20: Identification of genes with uORFs. A) Graphical representation of frequency distribution of nucleotides, which are found, at indicated positions relative to the ATG of 6692 main ORF translation start sites (sequences downloaded from Biomart. B) Screenshot of TL analysis where each row contains details about a uAUG identified in yeast. Columns from left to right are: Systematic Name, Gene Name, uORF Kozak score, frame (uAUGs that are in the same reading frame are labeled as 'IN' and those in a different reading frame are labeled as 'OUT'), length of uORF (if 0 then there is no uORF associated with the TL), TL\_length (the length of the TL), ORF Kozak score (the Kozak score associated with the initiation codon from the main ORF), length of oORF (if 0 then there is no oORF associated with the TL). The data displayed here is generated from sequences retrieved from SGD (<http://yeastmine.yeastgenome.org/yeastmine/service>) on 26.06.2017. Python scripts used to generate data can be found at (<https://github.com/vickytorrance/uAUGs>) C) Summary of B showing total uAUGs, uORFs and oORFs that were identified.

## **4.2 Genes with uORFs have lower translation rates and uORFs are underrepresented in the genome.**

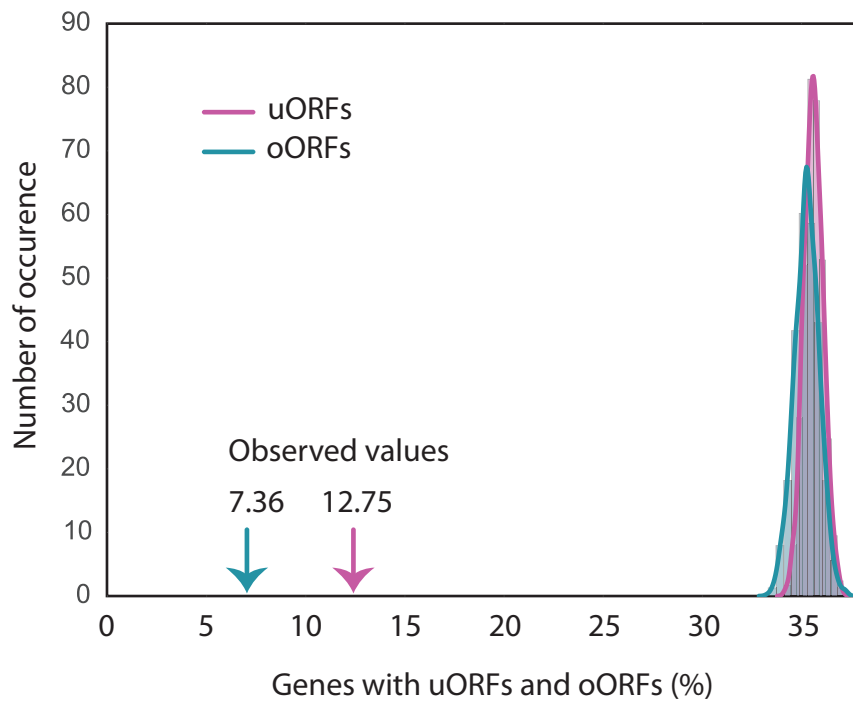
Since uORFs and oORFs are generally regarded as elements that inhibit translation we hypothesised that their occurrence would be lower than predicted by chance, reflecting negative selection. If oORFs inhibit expression more than uORFs we would expect selection against oORFs to be stronger than selection against uORFs. To examine the selective pressure on uORFs, random DNA sequences were generated to replace actual TL sequences and the percentage of those TLs that contained uORFs and oORFs were calculated. 1000 trials were performed, each trial generating a new random gene-specific TL sequence and calculating the percentage of those TLs that contained uORFs and oORFs. While 12.75% of genes contain uORFs and 7.36% contain oORFs the average number of uORFs and oORFs observed in randomly generated sequences is about 35% demonstrating strong selective pressure against uORFs and oORF (Figure 21a). The selective pressure against oORFs is greater than that against uORFs indicating that they may be stronger repressors of translation. The observation that there are less uORFs and oORFs than expected from random sequences is in agreement with published data showing that there are less uAUGs than expected (Arribere and Gilbert, 2013).

To further explore the effect of uORFs and oORFs on gene expression we compared estimated translation rates of genes containing either uORFs or oORFs with translation rates from all genes using a published dataset of estimated translation rates (Arava et al., 2003). As expected, the genes that contain uORFs or oORFs have reduced translation rates when compared with the full set of genes (Figure 21b). Interestingly genes with oORFs have slightly lower translation rates than the genes with uORFs supporting the hypothesis that oORFs inhibit translation to a greater degree than uORFs (Figure 21b).

There are several explanations as to why oORFs may be more repressive than uORFs. The out of frame translation events that may occur on transcripts that have oORFs may result in these being degraded by NMD (Celik et al., 2017a). Another explanation would be that uORFs permit translation re-initiation, which is not possible after translation of an oORF. oORFs in the same reading frame as the main ORF result in the synthesis of proteins with N terminal extensions that may negatively affect function. It is unsurprising therefore that of the genes that contain 464 oORFs,

only 46 have oORFs in the same reading frame as the main ORF, much lower than the expected 1 third that would occur if the oORFs were evenly distributed in each reading frame.

A



B

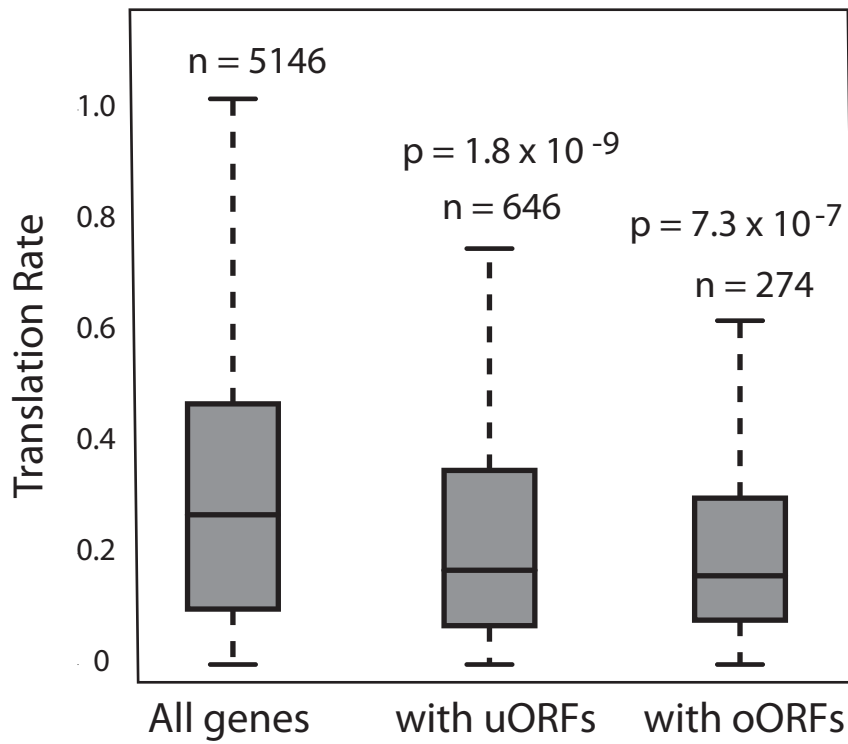


Figure 21: Genes with uORFs have lower translation rates and uORFs are underrepresented in the genome. A) The percentage of genes, which contain either uORFs or oORFs, and histogram of the percentage of uORF and oORF containing genes calculated from 1000 randomizations of gene specific TL sequences. B) Boxplot of the translation rates of all genes, genes with uORFs and genes with oORFs. Estimates of the translation rates are taken from (Arava et al., 2003). Authors estimated translation rate by multiplication of the fraction of transcripts engaged in translation with the density of ribosomes on a transcript. The box indicates location of the 25th and 75th percentiles, with the difference between these two percentiles being the interquartile range (IQR). The upper whisker indicates the largest data point less than  $1.5 * \text{IQR}$ . Similarly the lower whisker indicates the smallest data point greater than  $1.5 * \text{IQR}$ . Any data points outside of this range are considered to be outliers (not indicated). P values are calculated using unpaired t-test.

### 4.3 uORFs have weaker initiation contexts than main ORFs

The fact that uORFs and oORFs are heavily selected against and are associated with lower translation rates suggests that those uORFs that are present may play important regulatory roles, such as the *GCN4* uORFs discussed in the introduction (Figure 8). The nucleotide sequence flanking an ATG determines the likelihood of it being 'recognized' as an initiation codon (Kozak, 1986b), therefore if uORFs and oORFs do play an important regulatory role one might expect them to be in favorable initiation contexts. On the other hand if uORFs and oORFs do not play important regulatory roles we might expect unfavorable initiation contexts to be common.

To explore the biological relevance of uORF and oORF initiation contexts uORF and oORF Kozak scores were compared with main ORF Kozak scores and Kozak scores generated from random TL specific sequences, termed 'scrambled sequence'. The 'scrambled' TL sequences were constructed as described in 4.2. As expected Kozak scores from the main ORFs are higher than those generated from random sequences however Kozak scores from uORFs and oORFs were not statistically different from Kozak scores of randomly generated sequences suggesting that there is no general selective pressure acting on uORF initiation contexts (Figure 22). This is consistent with data from the *Chlamydomonas* genome where the Kozak scores of uORFs and randomly generated sequences are similar (Cross, 2015) and a study of mouse uORFs, which found no distinct sequence motif around the uORF initiation codon (Chew et al., 2016).

The finding that there is no selective pressure acting on uORF initiation contexts suggests that uORF initiation contexts do not play a major role in regulation of translation.



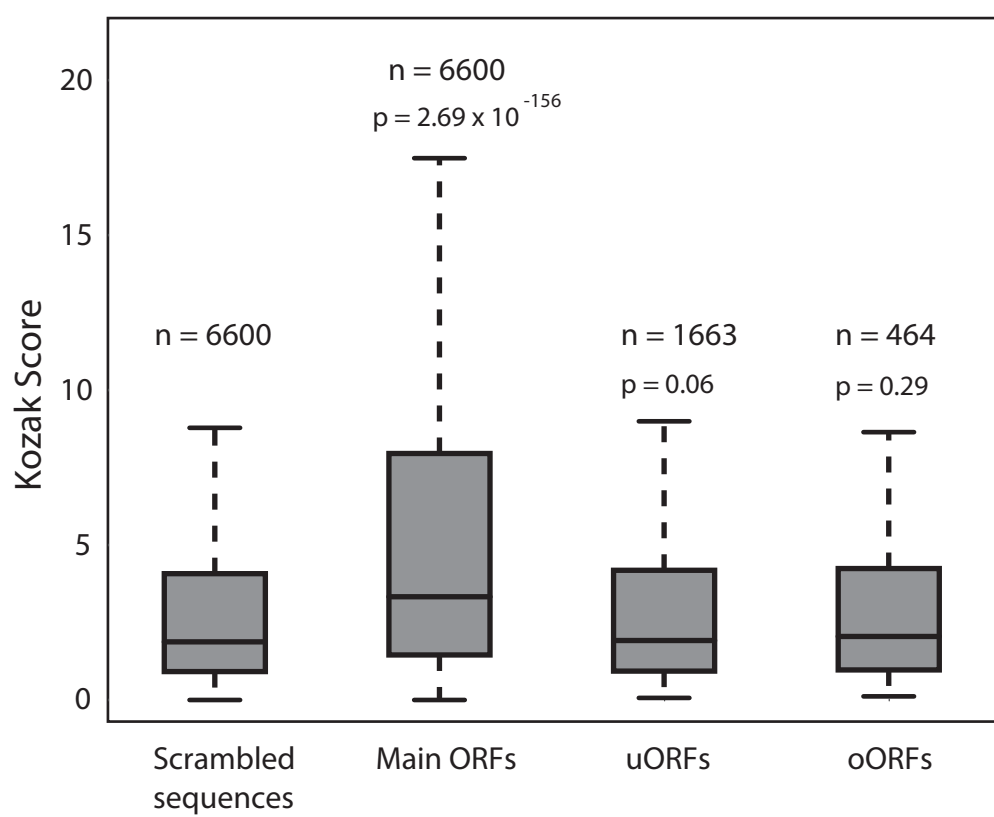


Figure 22: uORFs have weaker initiation contexts than the main ORFs. Boxplots of the all Kozak scores from uORFs, oORFs, main ORF as well as main ORFs generated from randomized DNA sequences (scrambled DNA). 'Kozak score' describe the initiation context surrounding initiation codon, by comparing likelihood to the consensus Kozak sequence. Description of how Kozak scores were calculated is in methods. Each boxplot represents the 'Kozak scores' calculated from each individual uORF, oORF or main ORF within the genome. As a comparison the 'Kozak scores' of initiation codons generated from a randomised DNA sequence are also presented. The box indicates location of the 25th and 75th percentiles, with the difference between these two percentiles being the interquartile range (IQR). The upper whisker indicates the largest data point less than  $1.5 * \text{IQR}$ . Similarly the lower whisker indicates the smallest data point greater than  $1.5 * \text{IQR}$ . Any data points outside of this range are considered to be outliers (not indicated). P values are calculated using unpaired t-test.

#### **4.4 uORFs are longer and terminate further from main ORFs than expected.**

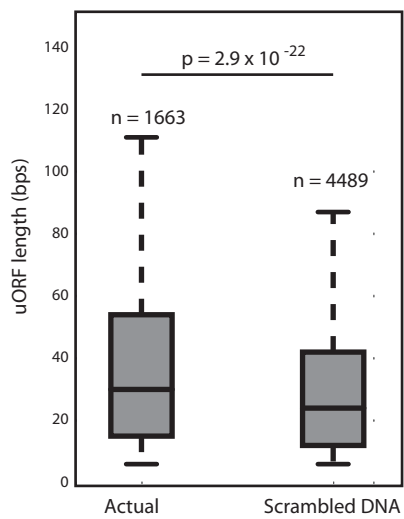
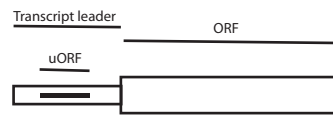
Since there seems to be no general selective pressure acting on the uORF initiation contexts we wondered if there is selection for other uORF features that may explain how they affect gene expression. The likelihood of translation re-initiation occurring is negatively correlated with uORF length and proximity to the main ORF (Rajkowitsch et al., 2004) (Hinnebusch, 2005). If translation re-initiation plays a major role in permitting downstream translation after a uORF then there may be selection for uORFs that are shorter and reside further from the main coding sequence.

Actual uORFs from the genome and uORFs generated from random TL specific sequences were compared in length and proximity to the main coding sequence. Distances between uORFs and corresponding coding sequences are significantly longer than expected from random sequences, indicating selection for uORFs that reside far from the main ORF (Figure 23a). This indicates that there may be selection for uORFs that facilitate translation re-initiation (Hinnebusch, 2005). However we also found that uORFs are significantly longer than expected from random sequences. Since longer uORFs reduce the likelihood of translation re-initiation occurring (Rajkowitsch et al., 2004) this conflicts with the idea that there is selective pressure acting on uORF features that facilitate translation re-initiation (Figure 23b).

One explanation for why there is selection in favour of uORFs that reside further from the main coding sequence is that these uORFs are more likely to be missing from some mRNA isoforms than uORFs that are in close proximity to the main coding sequence. The finding that there is not selection against longer uORFs, the products of which are presumably more energetically costly to synthesise compared shorter uORFs, might suggest that some uORFs encode functional peptides. A simpler explanation however is that shorter uORFs repress translation more than longer uORFs, possibly by recruiting NMD factors more efficiently or resulting in a greater number of stalled ribosomes that act as a 'roadblock' to translation as illustrated in Figure 7.

In summary, there is more selection against short uORFs and uORFs positioned close to the main ORF start codon leading to the hypothesis that these uORF features are more repressive to translation than longer uORFs or uORFs close to the m<sup>7</sup>G cap.

A



B

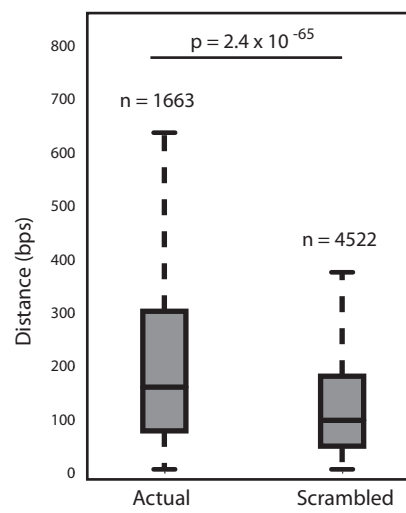
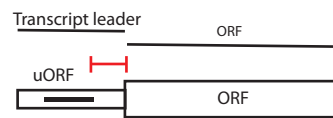


Figure 23: uORFs are longer and further from main ORFs than expected. A) Boxplots of uORF lengths (A) or distances between the uORF stop and main ORF (B) from uORFs found in actual TL sequences and uORFs generated from randomized gene specific TL sequences (scrambled DNA). The box indicates location of the 25th and 75th percentiles, with the difference between these two percentiles being the interquartile range (IQR). The upper whisker indicates the largest data point less than  $1.5 * \text{IQR}$ . Similarly the lower whisker indicates the smallest data point greater than  $1.5 * \text{IQR}$ . Any data points outside of this range are considered to be outliers (not indicated). P values are calculated using unpaired t-test.

#### 4.5 New luciferase reporters created to measure translation *in vivo*.

*TMA20* and *TMA22* homologs in *Drosophila* have been shown to promote expression of genes that contain uORFs, particularly those that contain short uORFs in strong initiation contexts, therefore we identified all of the genes in yeast that contain uORFs. If the function of Tma20 and Tma22 is conserved in yeast then such genes will be indicative of those regulated by Tma20 and Tma22. We therefore sought to develop an assay to test if Tma20 and Tma22 regulate gene expression and investigate the hypothesis that Tma20 and Tma22 promote translation re-initiation.

Plasmids were designed to quantify gene expression *in vivo* using two different reporter genes, Renilla luciferase and Firefly luciferase (Figure 24a). The gene encoding Firefly luciferase is fused to a constitutive promoter, *pPGK1*, to act as a 'loading control'. The gene encoding Renilla luciferase is fused with the test upstream regulatory sequence (URS), which encompasses the promoter and TL. To ensure that expression of the test URS is independent from the expression of the loading control, *pPGK1*, genes encoding Renilla luciferase and Firefly luciferase were fused to terminator sequences from *ADH1* and *CYC1* respectively and positioned in opposite orientations to each other. A dual luciferase assay (Promega) was used to measure Firefly luciferase (FLuc) and Renilla luciferase (RLuc) activity from a single sample and the ratio of RLuc/FLuc represents the 'normalised' expression of a test promoter.

For any given gene of interest the URS is amplified from genomic DNA using primers containing homology regions with the empty vector, to allow the promoter to be fused with Renilla luciferase. This plasmid could then be transformed into *WT* and *tma20Δ* strains to measure if indeed *TMA20* regulates expression of the URS.

To confirm that the regulation of the URS, by *TMA20*, is due to the uORF the effect of *TMA20* on expression of a control URS can also be measured using a control plasmid. The control plasmid is identical apart from a point mutation in the uORF/s initiation codon. Point mutations are introduced using a PCR based strategy (see Figure 24a).

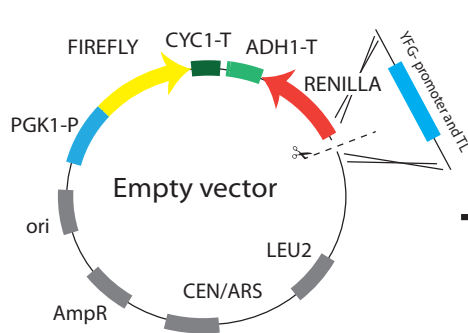
To test the efficiency of a test promoter, Firefly luciferase and Renilla luciferase activity were measured in cells transformed with a plasmid with or without (empty vector) an *ALD3* promoter fused to the Renilla luciferase gene. Although, as

expected, Firefly luciferase activity is detected from the empty vector and the vector with *ALD3* fused to Renilla luciferase, Renilla luciferase activity is only detected from the latter vector (Figure 24B).

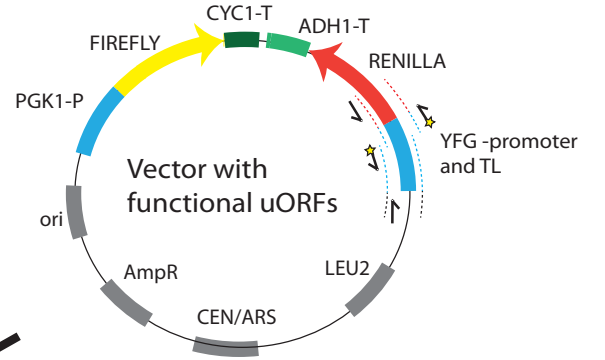
To test the dynamic range of the assay a 2 fold dilution series of cells transformed with reporter plasmid was made. Renilla luciferase was expressed under *HDA3* or *HDA3-M1V* (a variant that contains a point mutation in the USR of *HDA3*). As anticipated Firefly luciferase and Renilla luciferase activities increased linearly with increasing concentrations of cells, in both samples, despite the difference in Renilla luciferase expression detected from *HDA3* or *HDA3-M1V* (Figure 24c). We repeated the test in *tma20Δ*, also finding Firefly luciferase and Renilla luciferase activities to increase linearly with increasing concentrations of cells.

**A**

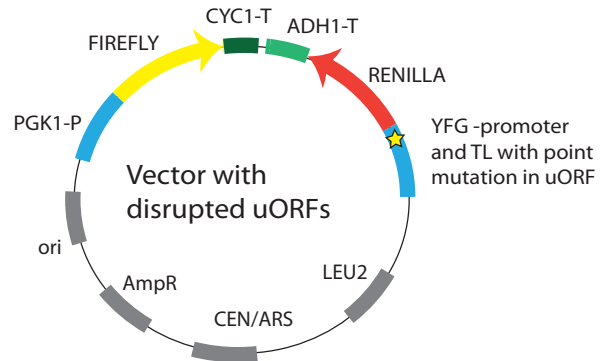
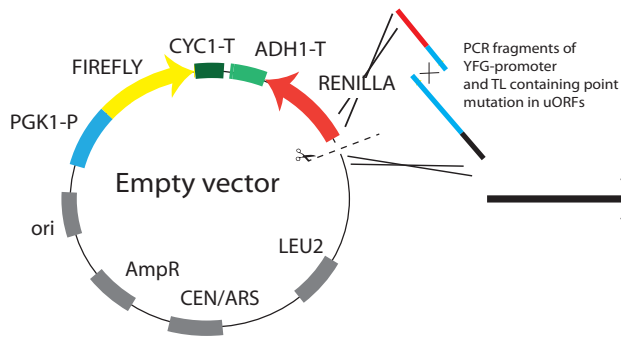
Step 1: Clone the promoter and TL into empty vector



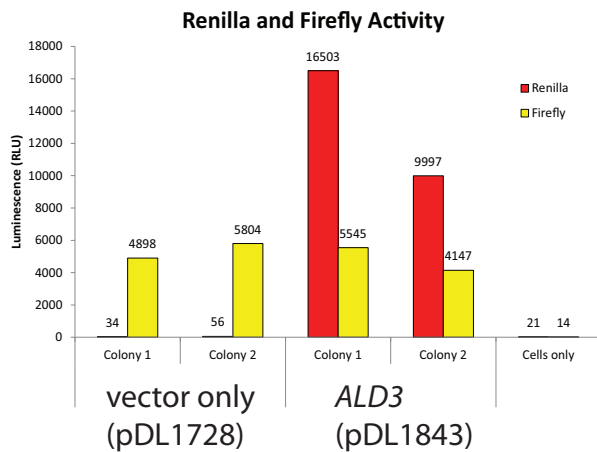
Step 2: PCR amplify the promoter and TL using primers containing point mutations



Step 3 Clone PCR fragments of promoter and TL containing point mutations



**B**



**C**

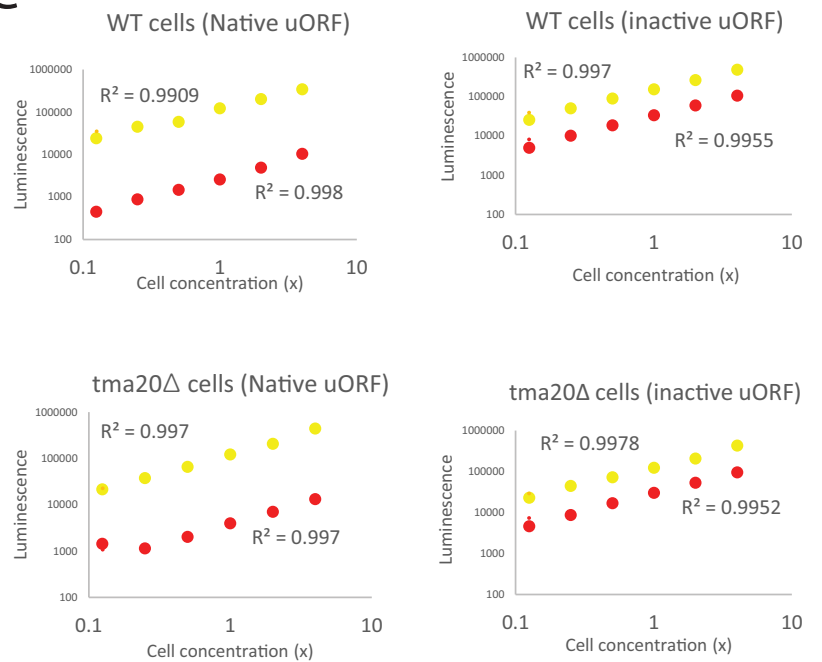




Figure 24: New luciferase reporters created to measure translation *in vivo*. A) Diagram illustrating design and method of assembly of reporter plasmids used in this study. *YFG* promoter drives Renilla luciferase and *PGK1* promoter drives Firefly luciferase, which is used as a loading control. Briefly, promoter of *YGF* is amplified from genomic DNA and cloned into 'empty' reporter plasmid. Point mutations are introduced onto primers, which amplify part of *YFG* promoter and a region of the plasmid. Two or more primer pairs are used to amplify two or more fragments of the promoter/plasmid DNA, which are subsequently cloned into the 'empty' reporter plasmid. B) Luminescence signals detected from cells transformed with plasmid with (pDL1843) and plasmid without (pDL1728) an *ALD3* promoter fused to Renilla luciferase. Cells were grown and luminescence detected according to the protocol detailed in 'Luciferase assay' section of the methods. C) Luminescence from Firefly luciferase and Renilla luciferase measured from varying concentrations of cells. Cells were grown to exponential phase and

#### 4.6 uORFs, as expected, inhibit the expression of main ORFs to varying extents.

To explore the role of Tma20 and Tma22 in translation, 22 genes were selected based on the criteria that they contain uORFs in strong initiation contexts (thus have higher Kozak scores) and have TL length predictions published in (Arribere and Gilbert, 2013) and (Nagalakshmi et al., 2008) that are roughly in agreement with each other (Figure 25).

The TL and promoter regions of the 22 genes that we selected were each fused to a test reporter gene (Renilla luciferase). The initiation codons of each of the uORFs were subsequently eliminated by point mutation, thereby creating pairs of identical plasmids, one containing a promoter and TL with native uORF/s and one containing a promoter and TL with disrupted uORF/s (Figure 24a). To visualise the uORFs and TLs of each of these genes, the promoter, TL and uORFs are drawn, to scale (Figure 25).

Before investigating the regulation of expression of these 22 genes by *TMA20* and *TMA22* we first considered the impact that uORFs have on expression of these 22 genes. The effect that uORFs have on expression is quantified using the ratio of normalised Renilla luciferase activity when uORFs have been eliminated relative to when uORFs have not been eliminated.

In summary of the 22 genes we examined, 13 have uORFs that repress expression, 7 have uORFs that do not affect expression and 2 have uORFs that seemingly promote expression. We grouped the 22 selected genes according how much expression changes upon elimination of the uORFs (Figure 27a). The first group (*CTF3*, *BRE4*, *HDA3*, *ATG5*, *NTR2*, *NDJ1*, *PSG1* and *FRE6*) have uORFs that greatly reduce expression, as demonstrated by a greater than 2 fold increase in normalised reporter activity upon elimination of uORF/s. Elimination of *CTF3* and *BRE4* uORFs results in the most dramatic increase in expression that is remarkably 16 fold and 12 fold respectively. The second group (*RRG1*, *PUS9*, *MOD5*, *ATG20* and *SMY1*) have uORFs that slightly reduce expression. We observed a less than 2 fold increase in normalised reporter activity that is statistically significant upon elimination of the uORFs within the second group. The third group of genes (*DUS3*, *MBP1*, *RBS1*, *SFG1*, *CDC13*, *IST2* and *RAI*) have uORF/s that do not have a

statistically significant impact on expression of the reporter gene. Finally eliminating the uORFs of the fourth group *CMK1* and *ALD3* unexpectedly resulted in a decrease in expression, suggesting that these uORFs may actually promote expression (Figure 26U-V).

We observed great variation in the influence that uORFs have on expression, therefore wondered if these variations might be attributed to particular uORF features. Since we observed that uORFs tend to be longer and reside further from the main coding sequence than expected from random TL sequences (Figure 23) we speculated that these uORFs would repress translation less than shorter uORFs that are in close proximity to the main ORF.

As anticipated we found that, of the 8 genes whose expression increases by more than 2 fold upon elimination of the uORFs (group 1), 7 of them have a uORF that terminates close to the main coding sequence (Figure 27b). In contrast, of the 7 genes that have uORFs that do not influence gene expression (group 3), only 1 has a uORF that terminates close to the main coding sequence (Figure 27b). However, in contrast with our observation that short uORFs are less common than expected by random sequences (Figure 23a) we found no obvious connection between the degree to which a gene is repressed by a uORF/s the presence of a short uORF (Figure 27c). We also found no evidence to suggest that long uORFs correlate with strong suppression of gene expression (Figure 27d). There is no correlation between the initiation context of the uORFs within a TL and the change in gene expression observed upon elimination of the uORFs (Figure 27e), consistent with our previous observation that there is no deviation in the uORF Kozak scores from those expected by random sequences (Figure 22). This analysis is somewhat complicated because some TL's having multiple uORFs however valid comparisons can still be made, in that we can ask whether having, for example a short uORF correlates with strong uORF dependant regulation of gene expression.

Since mRNAs frequently exist in multiple isoforms we predict that, in some genes, there will be transcript isoforms that do not contain the uORF so the overall influence that the uORF has on expression will be reduced. In support of this idea, it was recently suggested that about 70% of genes transcribe multiple isoforms of mRNA (Pelechano et al., 2013).

The decrease in *ALD3* and *CMK1* expression observed upon introduction of a point mutation to their respective uORFs indicates that expression of *ALD3* and *CMK1* is facilitated by uORFs, in contradiction with the general view that uORFs impede translation (Jackson et al., 2010). An alternative explanation however is that the point mutations we introduced affect secondary structure of the TL, thereby negatively influencing translation. It would be of interest to determine a mechanism by which uORFs can promote downstream expression.

Tma20 and Tma22 are thought to be required for optimal translation of genes that have uORFs. Since we demonstrated that uORFs have different effects on translation we speculated that deletion of *TMA20* and *TMA22* would cause a reduction expression of those genes whose expression is affected by uORFs.

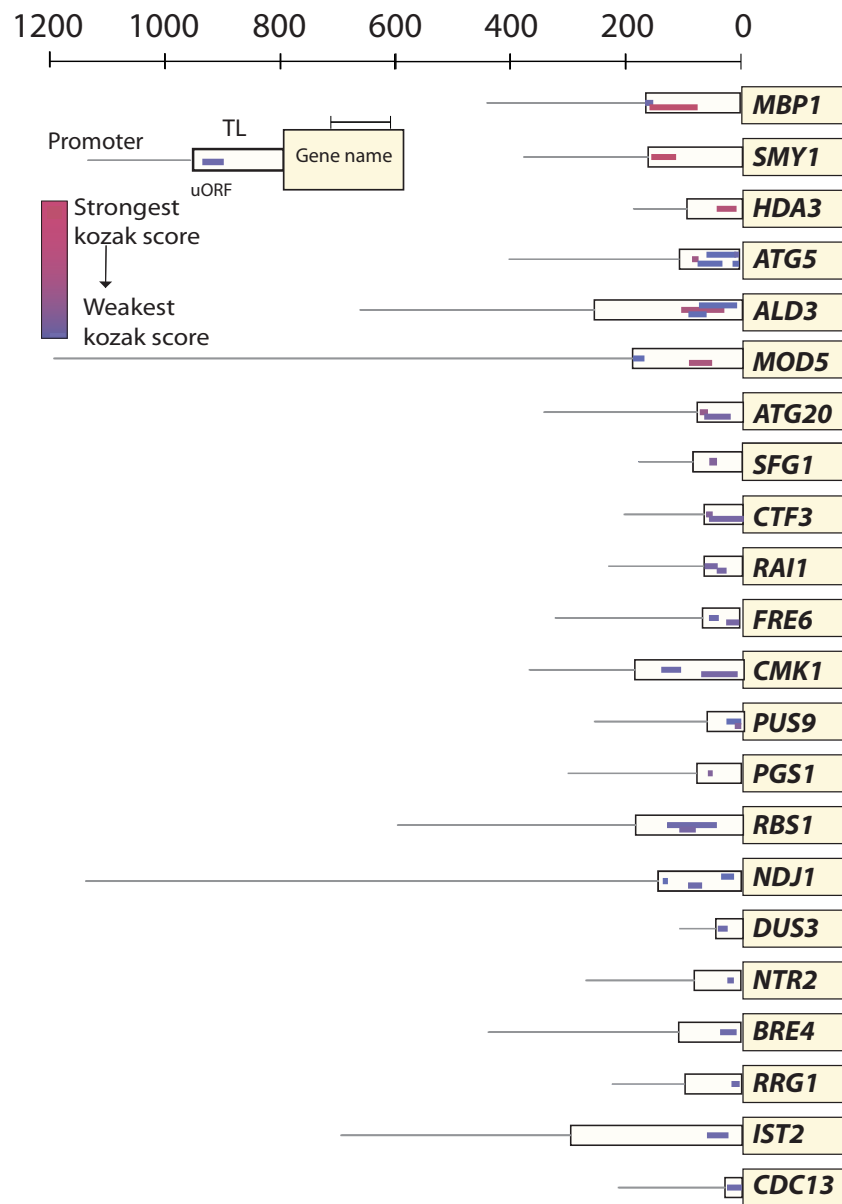
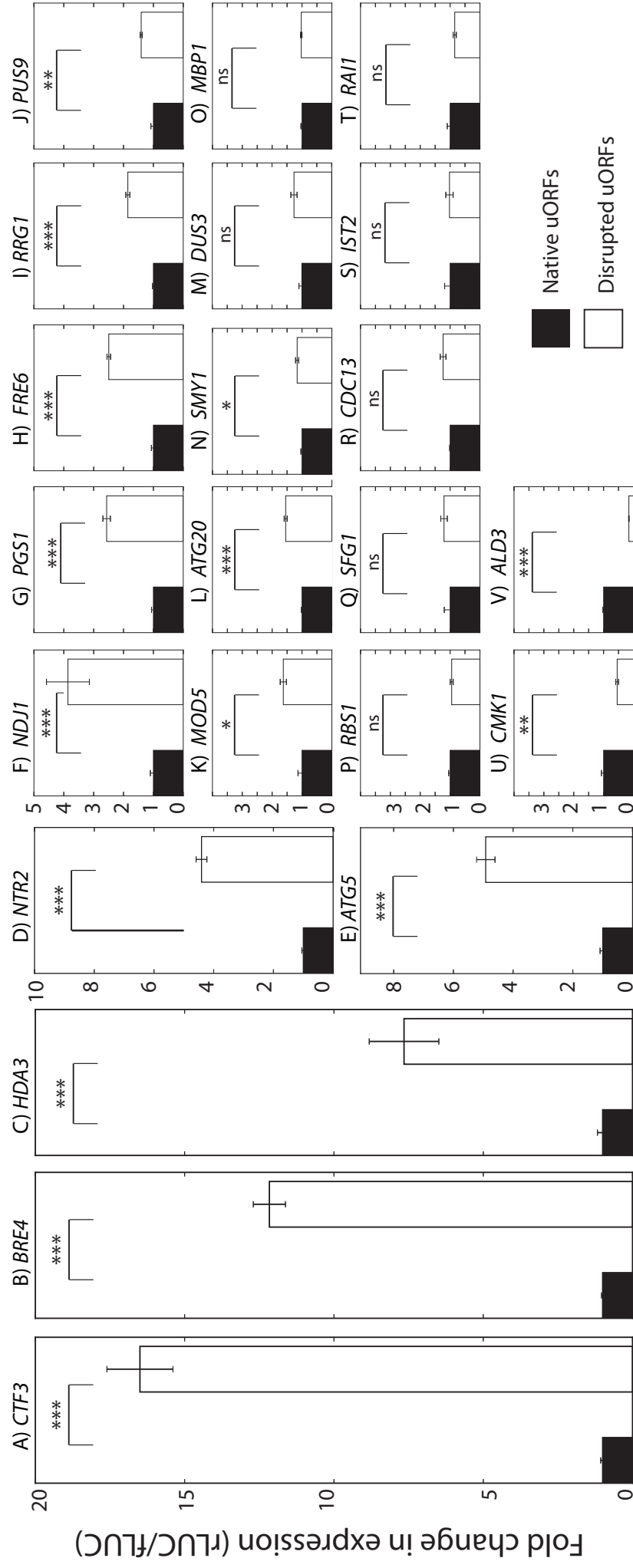


Figure 25: 22 genes with uORFs selected to explore the role of Tma20 and Tma22 in translation. Illustration depicting the genes used to measure effect of mutating uORFs. Promoter lengths were taken from Yeast Promoter Atlas (<http://archive.is/Xgoi>) accessed on 05/12/2015 (Chang et al., 2011), TL length estimates taken from (Arribere and Gilbert, 2013) and (Nagalakshmi et al., 2008). Colours have been used to indicate the relative strength of the Kozak sequence.



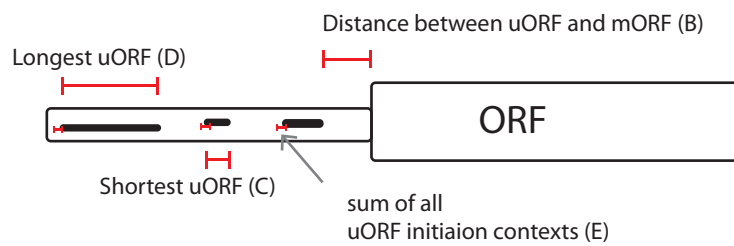
**Figure 26 - uORFs, as expected, inhibit the expression of main ORFs to varying extents.**

Figure 26: uORFs, as expected, inhibit expression of main ORFs to varying extents. A-U shows the effect of abolishing uORFs (by introducing a point mutation in the initiation codon). *YFG* expression is determined as the ratio of Renilla (*YFG* driven) over Firefly (*PGK1* driven, normalisation control) luciferase activity (*rluc/fluc*). The effect of abolishing uORFs is quantified by expressing normalised Renilla activity (*rluc/fluc*) when uORFs have been abolished (unfilled bars) relative to normalised Renilla activity (*rluc/fluc*) when uORFs have not been abolished (filled bars). Cells were grown according to 'Luciferase assay' protocol in the methods section. Plasmids were transformed into two independent WT strains and luciferase activity was measured in two colonies from each transformation reaction. Means of four independent repeats are plotted with error bar indicating the SEM. *P* values were calculated using t-test (\*)  $P < 0.05$ , (\*\*)  $P < 0.01$ , (\*\*\*)  $P < 0.001$ .

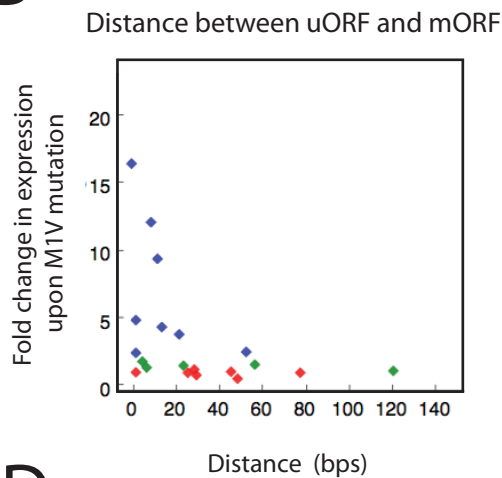


A

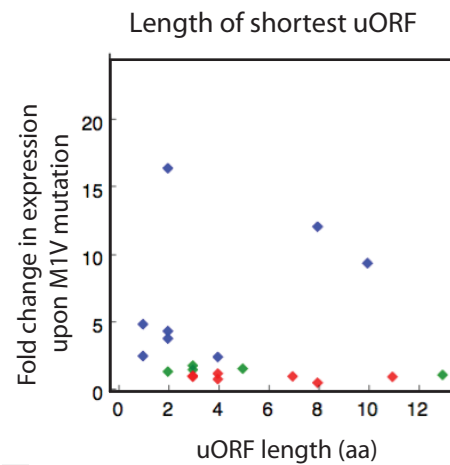
Change in gene expression that occurs upon introducing point mutations into uORFs		Gene
Group 1 <span style="color: blue;">◆</span>	Increases more than 2 fold	<i>CTF3, BRE4, HDA3, ATG5, NTR2, NDJ1, PGS1, FRE6</i>
Group 2 <span style="color: green;">◆</span>	Increases between 1 and 2 fold	<i>RRG1, PUS9, MOD5, ATG20, SMY1</i>
Group 3 <span style="color: red;">◆</span>	No significant change	<i>DUS3, MBP1, RBS1, SFG1, CDC13, IST2, RAI1</i>
Group 4	Decrease	<i>CMK1, ALD3</i>



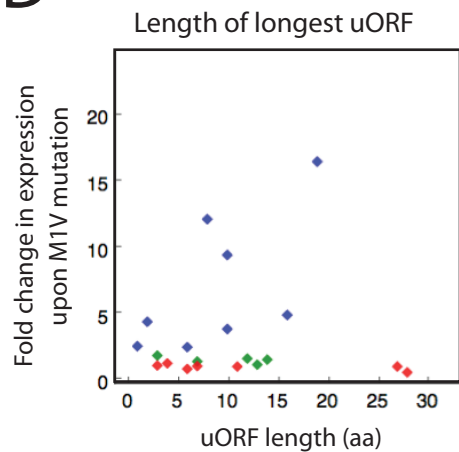
B



C



D



E

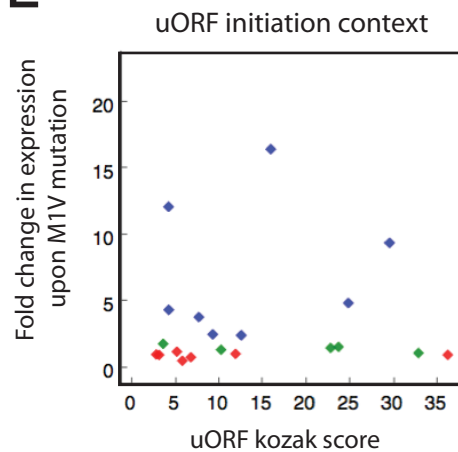


Figure 27- uORFs that reside close to the main ORF correlate with reduced translation. A) Summary of the impact that eliminating the uORFs of different transcripts has on their expression. Each gene is grouped according to the change in expression that occurs when uORFs are eliminated from the transcript. B-E) Comparison of the change in gene expression that occurs when uORFs have been eliminated from a transcript with B) the distance separating the uORF termination codon and the main ORF initiation codon, C) the length of the shortest or D) longest uORF within the TL, E) the initiation contexts of the uORFs. The colours refer to the particular grouping of each gene that is shown in A.

#### 4.7 No evidence that Tma20 or Tma22 promote translation re-initiation in yeast

To test if the 22 candidate genes selected in this study are regulated by *TMA20* and *TMA22* we measured the effect of *tma20Δ*, *tma22Δ* and *tma20Δ tma22Δ* on each of the selected genes. To clarify that the effect of *TMA20* and *TMA22* deletion is indeed because of the uORF, the effect of *tma20Δ* and *tma22Δ* on expression of identical reporter harbouring point mutations in the uORF initiation codons was also measured. If Tma20/Tma22 complex promotes translation re-initiation we expect *tma20Δ* and *tma22Δ* to result in a decrease in expression since Tma20/Tma22 complex would be required for efficient re-initiation following translation of a uORF. However, when the uORFs have been disrupted, we expect the deletion of *TMA20* or *TMA22* to have no effect. Since Tma20 and Tma22 function as a heterodimer (Lomakin et al., 2017) we predict that *tma20Δ*, *tma22Δ* and *tma20Δ tma22Δ* will function as biological controls. The mammalian homolog of Tma64, eIF2D has been shown to share functions with the heterodimer of Tma20 and Tma22. Both eIF2D and Tma20/Tma22 complex promote the dissociation of 40S, mRNA and de-acetylated tRNA following translation termination (Skabkin et al., 2010a) and can also facilitate re-initiation events on viral transcripts (Zinoviev et al., 2015). We therefore investigated the effect of *tma64Δ* on expression of some of the 22 candidate genes selected.

In 10 out of 22 genes (*MBP1*, *MOD5*, *SMY1*, *NTR2*, *DUS3*, *PUS9*, *RRG1*, *BRE4*, *CTF3* and *ALD3*) expression was higher in *tma20Δ*, *tma22Δ* and *tma20Δtma22Δ* strains compared with WT or *tma64Δ* strains suggesting that Tma20/Tma22 complex but not Tma64 inhibits the expression of these genes (Figure 28 – indicated with diamonds). However, we also observe, in 9 of these 10 genes, that *tma20Δ*, *tma22Δ* and *tma20Δ tma22Δ* result in a similar increase in expression when uORF initiation codons have been eliminated demonstrating that the effect of *TMA20* and *TMA22* is independent of uORFs (Figure 28 –indicated with filled diamond). In just one gene, *BRE4*, *tma20Δ* and *tma22Δ* increase expression when the uORF is intact and not when it had been inactivated, consistent with a model that Tma20 and Tma22 repress rather than promote, translation re-initiation (Figure 28 – indicated with filled diamond). In the other 12 genes investigated expression is similar in *tma20Δ*,

*tma22Δ*, *tma20Δtma22Δ*, *tma64Δ* and wild type strains, both when uORFs were intact and when they had been disrupted suggesting that *TMA20*, *TMA22* or *TMA64* do not regulate expression of these genes.

Interestingly the impact of *tma64Δ* on expression of most genes was distinct from the effect of *tma20Δ*, *tma22Δ* and *tma20Δ tma22Δ*, in agreement our observation that *TMA20* and *TMA22* function differently to *TMA64* at telomeres. This provides further evidence that *TMA20/TMA22* and *TMA64* have independent functions in yeast.

Our data demonstrates that, in contrast with their *Drosophila* homologs, *TMA20* or *TMA22* are not required for the efficient translation of genes that contain uORFs. Further we found evidence that Tma20 and Tma22 may actually act to repress expression after translation of some uORFs since they decrease expression of the *BRE4* allele that has a uORF but not the *BRE4* allele that does not have a uORF. Further experiments are required to confirm that the decrease in expression of *BRE4* promoted by Tma20 and Tma22 is actually because Tma20 and Tma22 repress translation re-initiation, rather than start codon recognition or transcript stability. Although our model that *TMA20* and *TMA22* repress translation re-initiation conflicts with the model in *Drosophila*, that they promote translation re-initiation, evidence that they promote dissociation of mRNA, tRNA and 40S following translation termination ribosome complexes are support the idea that Tma20 and Tma22 inhibit translation re-initiation. Dissociation of the 40S subunit from mRNA would prevent any possible re-initiation event from occurring.

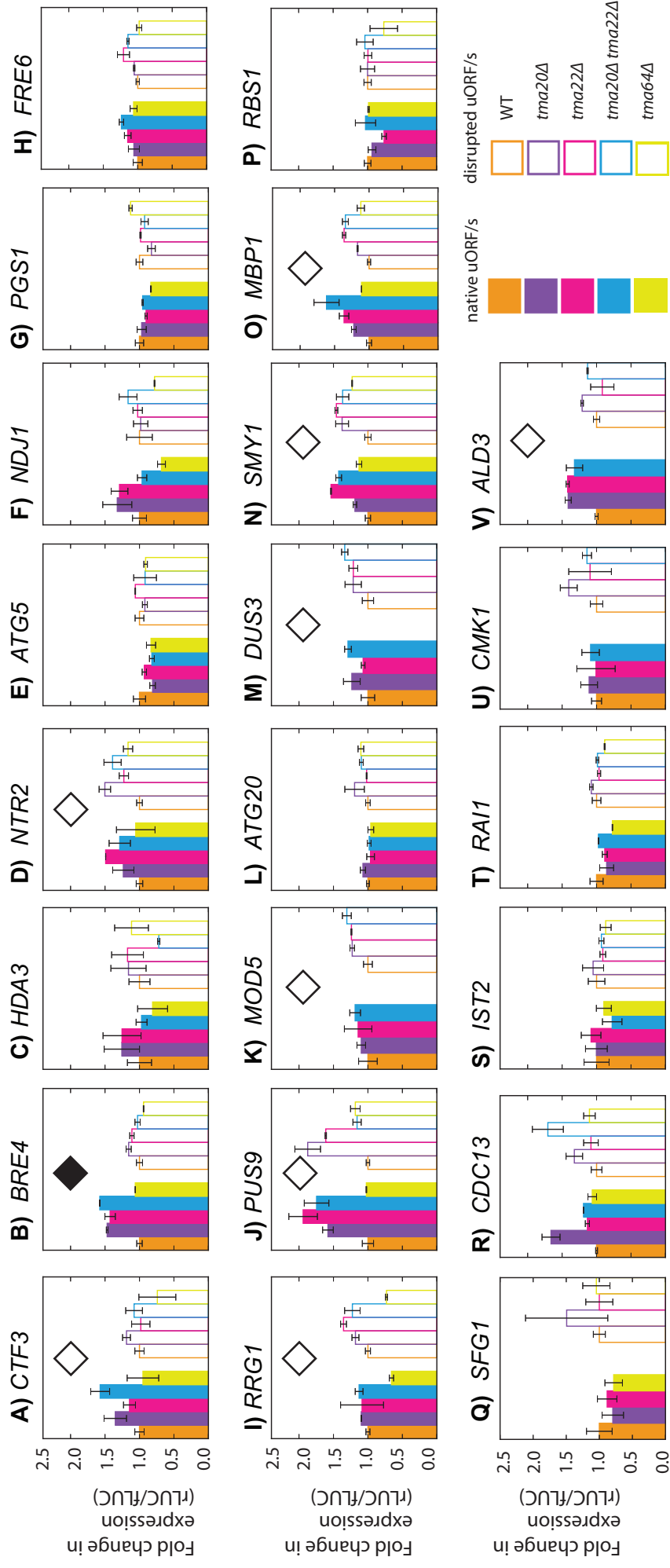


Figure 28. No evidence that Tma20 or Tma22 promote translation re-initiation in yeast. A-U) shows the effect, relative to *WT* of *tma20Δ*, *tma22Δ*, *tma20Δ tma22Δ* and in some cases *tma64Δ* on expression of two alleles of *YFG* (TLs containing uORFs are unfilled bars; mutated TLs with no uORFs are filled bars). The ratio of Renilla (*YFG* driven) over Firefly (*PGK1* driven, normalisation control) luciferase activity is calculated to obtain the normalised Renilla luciferase activity (*rluc/fluc*). For each plasmid pair (native uORFs; filler bars and disrupted uORFs; unfilled bars) the normalised Renilla luciferase activity of the *WT* was given a value of 1 and all other strains were expressed relative to this value. In deletion strains normalised Renilla luciferase activity was measured in two independent colonies and the bar indicates the mean of the two values, with the error bars indicating the individual values. In *WT* strains normalised Renilla luciferase activity was measured in four independent strains the bar indicates the mean of the four values with the error bars indicating the SEM of the four values. Cells were grown and luciferase assays conducted as per the conditions that are outlined in the ‘Luciferase assay’ section of the methods.

#### 4.8 Evidence suggests that Tma20 and Tma22 decreases transcript stability, as demonstrated using alleles of *BRE4*

*BRE4* is the only gene; out of the 22 we tested, for which *TMA20* and *TMA22* reduce expression when the uORF is present but not when the uORF is absent. This suggests that, in contrast with the proposed function of Tma20 and Tma22 homologs in *Drosophila* they inhibit rather than promote translation downstream of a uORF and do so in a gene specific rather than global manner. We therefore wanted gain further insight into the mechanism by which Tma20 and Tma22 inhibit expression of *BRE4*.

*BRE4* has a single 9 amino acid long uORF, situated in the same reading frame and just 9 nucleotides upstream of the main coding sequence (Figure 29a). To examine if Tma20 and Tma22 inhibit *BRE4* expression via the uORF initiation or termination codons we created an allele of *BRE4*, termed *BRE4-uORF-9\*Y* where the uORF termination codon is abolished by point mutation (Figure 29a). Since the uORF and main coding sequence are in the same reading frame *BRE4-uORF-9\*Y* encodes an isoform of Bre4 (or Renilla luciferase in this case) containing a 12 amino acid N-terminal extension. Therefore, to distinguish between stop codon read-through and translation re-initiation we also created an additional *BRE4* allele where the uORF and main coding sequence are in different reading frames, termed *BRE4-uORF out*. *BRE4-uORF-out* was created by insertion of a nucleotide 5' of the uORF stop codon resulting in a uORF that terminates 14 nucleotides downstream of *BRE4* main coding sequence. We also measured the effect that each mutation has on transcript levels to distinguish between effects on translation and on the stability of the transcript.

*BRE4-uORF-M1V*, which contains a point mutation in the *BRE4* uORF initiation codon, results in 12 fold increase in reporter protein activity, in agreement with previous data (Figure 29b). The 12 fold increase in protein expression is accompanied only by 2 fold increase in mRNA level indicating that the uORF represses expression mostly by decreasing translation, but in combination with reducing transcript levels (Figure 29 b+c). *BRE4-uORF-9\*Y* results in a more modest, 2 fold increase in reporter protein activity, despite the fact that, like *BRE4-uORF-M1V*, this allele does not have a uORF indicating that the N-terminal extension in *BRE4-uORF-9\*Y* may reduce the function of Renilla luciferase (Figure 29b). *BRE4-uORF-9\*Y* also results in a 2 fold increase in mRNA levels, suggesting that the increase in protein might be due to an increase in transcript level and not from

activity of Renilla luciferase that contains an N-terminal extension (Figure 29c). Interestingly expression of *BRE4-uORF-out* is reduced by about 25% compared with expression of *BRE4* showing that the overlapping version of the uORF represses *BRE4* expression more than the native uORF (Figure 29b). One possible explanation for this could be that *BRE4* expression is partially facilitated by translation re-initiation and translation re-initiation does not occur in *BRE4-uORF-out* since the termination codon resides within the main ORF.



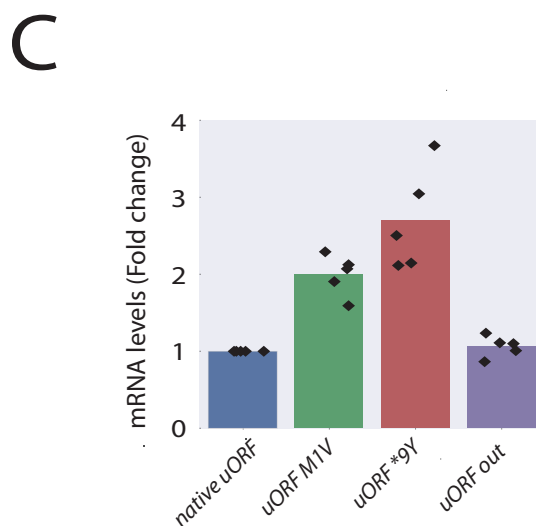
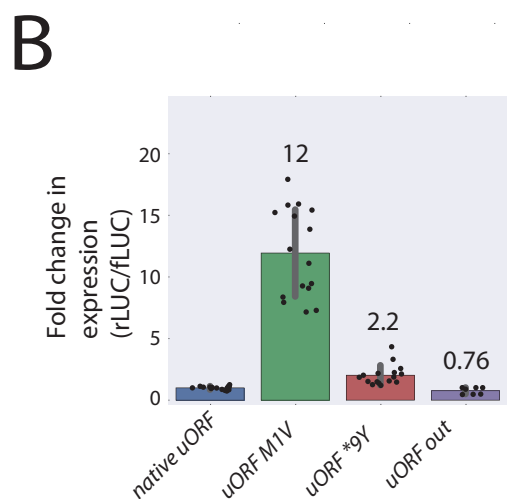
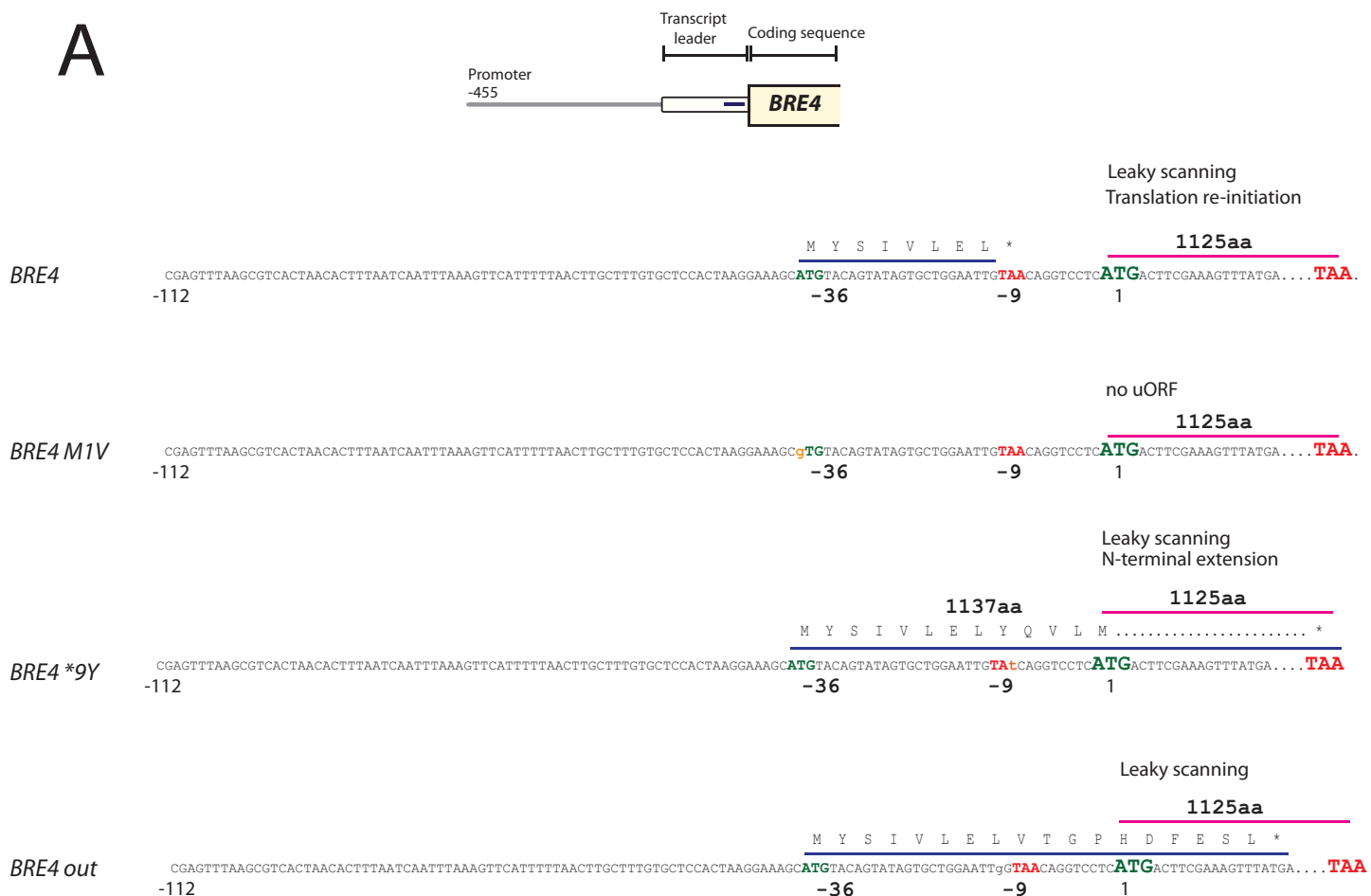


Figure 29. *BRE4* expression is decreased by a uORF. A) The TL sequence of *BRE4*, *BRE4-uORF-M1V*, *BRE4 uORF \*9Y* and *BRE4-uORF-out* indicating the positions of the uORF. *BRE4* has a single 9 amino acid long uORF. *BRE4-uORF-M1V* contains a point mutation in the uORF initiation codon and *BRE4-uORF-9\*Y* contains a point mutation in the uORF termination codon. *BRE4-uORF-out* was created by insertion of a nucleotide 5' of the uORF stop codon resulting in a uORF that terminates 14 nucleotides downstream of *BRE4* main coding sequence. B) The effect of each allele in A is quantified by expressing normalised Renilla luciferase activity (*rluc/fluc*) of *BRE4 uORF M1V*, *BRE4 uORF \*9Y* or *BRE4-uORF-out* relative to normalised Renilla luciferase activity (*rluc/fluc*) of *BRE4*. Each point represents an individual measurement (from an independent colony of cells), coloured according to when the experiment was conducted. Plasmid were transformed into WT (DLY 3001 or DLY8460) strains. Cells were grown and luciferase assay was conducted according to 'Luciferase assay' section of the methods. C) Transcript levels of Renilla luciferase measured in *BRE4*, *BRE4 uORF M1V*, *BRE4 uORF \*9Y* and *BRE4-uORF-out* measured was measured. RNA was isolated from cells grown in liquid culture to exponential phase at 30°C and transcript levels determined by SYBR Green RT-PCR. RNA concentrations of the samples were normalized to the loading control *BUD6*. The level of normalized Renilla luciferase expressed from *BRE4* was given the value of 1 and all other values were corrected relative to this. Each measurement was performed in triplicate. Each point on the plot represents an independent measurements (mean of triplicates) and the bar represents the mean of the 5 independent measurements.

To explore the function of Tma20 and Tma22 we next investigated the effect of *tma20Δ* and *tma22Δ* on expression of *BRE4*, *BRE4-uORF-M1V*, *BRE4-uORF \*9Y* and *BRE4-uORF-out*. As previously demonstrated *tma20Δ* and *tma22Δ* increase expression of *BRE4* but have no effect on expression of *BRE4 uORF-M1V* demonstrating that Tma20 and Tma22 repress *BRE4* expression by a mechanism dependant on the uORF (Figure 30a). In further support of the idea that regulation of Bre4 by Tma20 is dependant on the uORF *tma20Δ* increases expression of *BRE4-uORF-out*, a *BRE4* allele that contains an overlapping uORF, but has no effect on the expression of *BRE4 uORF \*9Y*, an allele which does not contain a uORF (Figure 30a).

Drosophila homologs of Tma20 and Tma22 are thought to promote expression of genes that contain uORFs by promoting translation re-initiation. The finding that Tma20 and Tma22 decrease expression of *BRE4* but not *BRE4-uORF-M1V* or *BRE4-uORF \*9Y* are in support of a model that Tma20 and Tma22 inhibit translation re-initiation. However Tma20 and Tma22 also decrease translation of *BRE4-uORF-out*, which presumably would not be permissive of translation re-initiation since it overlaps with main coding sequence (Figure 30a). This led us to consider other mechanisms by which Tma20 and Tma22 might repress *BRE4* expression.

Since levels of Renilla luciferase are lower when expressed from alleles of *BRE4* that contain uORFs this made us speculate that Tma20 and Tma22 may decrease levels of transcripts that contain a uORF. To test this we measured the effect of *tma20Δ* on reporter mRNA levels in each of the alleles of *BRE4*. Consistent with our hypothesis we observed that the increase in level of *BRE4* and *BRE4-uORF-out* reporter activity observed upon deletion of *TMA20* is actually accompanied by a similar increase in the mRNA level suggesting that *TMA20* decreases *BRE4* transcript levels rather than translation (Figure 30c).

Since transcripts that contain uORFs are NMD substrates (Gaba et al., 2005a) and we previously demonstrated that *TMA20* and *NMD2* act in the same pathway to suppress *cdc13-1* (Figure 15) we speculated that *TMA20* might repress *BRE4* and *BRE4-uORF-out* through NMD. Surprisingly however our data suggest that Tma20 represses expression independently of NMD since expression of *BRE4-uORF-out* is higher in *tma20Δ nmd2Δ* compared with *tma20Δ* or *nmd2Δ* (Figure 30b + c).

In summary, Tma20 and Tma22 repress expression of *BRE4* by reducing transcript levels and the mechanism by which they reduce transcript levels is via the uORF but is independent of NMD. Tma20 and Tma22 homologs in *Drosophila*, on the other hand, increase expression of transcripts that contain uORFs by increasing translation re-initiation. We propose a model where interaction of Tma20/Tma22 with the post termination complex can have multiple consequences to expression. In some cases, as with the *BRE4* example, Tma20/Tma22 decrease the level of transcript, presumably by decreasing its stability. In other cases, such as has been proposed in *Drosophila* (Schleich et al., 2014b), Tma20/Tma22 increase the efficiency of translation re-initiation. Our proposed model that Tma20/Tma22 interactions with uORFs can result in more than one outcome for expression is supported by biochemical evidence that MCT-1/DENR, like most initiation factors, adopt multiple roles in translation (Skabkin et al., 2010a). MCT-1/DENR promotes the recruitment of Met-tRNA<sub>i</sub><sup>Met</sup> to and also dissociation of 40s/tRNA/mRNA after translation termination (Skabkin et al., 2010a).

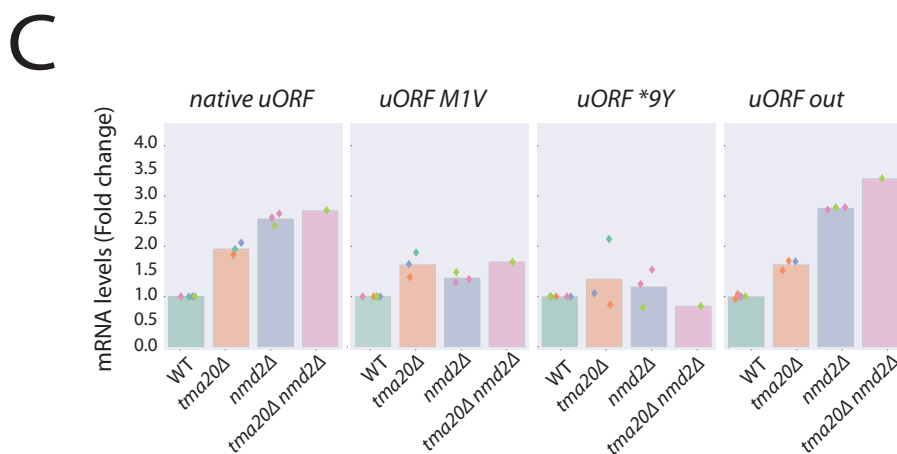
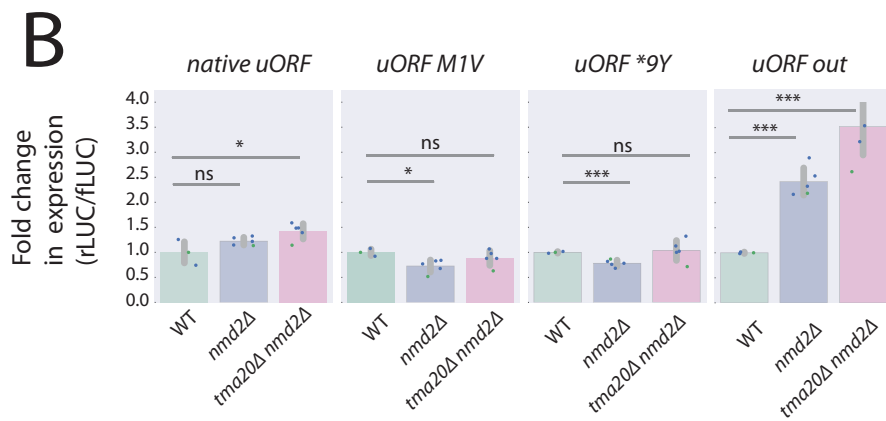
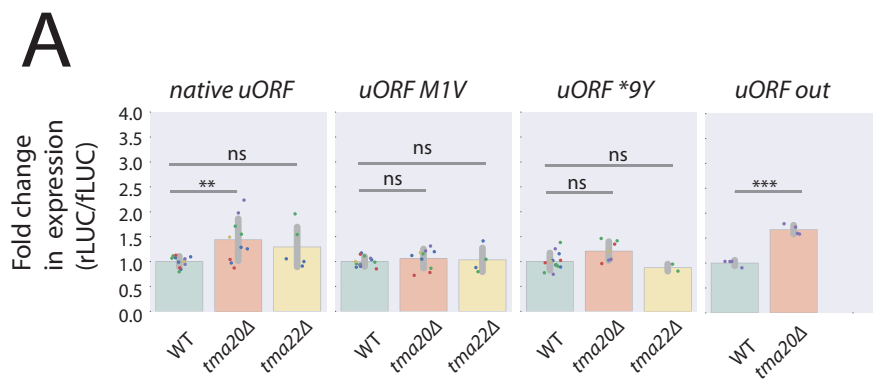


Figure 30. Tma20 and Tma22 decrease expression of *BRE4* by decreasing transcript levels. A) The effect, relative to *WT* of *tma20Δ*, *tma22Δ*, B) *nmd2Δ* and *tma20Δ nmd2Δ* on expression of *BRE4*, *BRE4-uORF-M1V*, *BRE4-uORF-\*9Y* and *BRE4-uORF-out*. The ratio of Renilla (*BRE4* driven) over Firefly (*PGK1* driven, normalisation control) luciferase activity (*rluc/fluc*) is calculated. Each point represents an individual measurement, coloured according to when the experiment was conducted. Cells were grown and luciferase assays conducted as per the conditions that are outlined in the 'Luciferase assay' section of the methods. C) Transcript levels of Renilla luciferase expressed from *BRE4*, *BRE4-uORF-M1V*, *BRE4-uORF-\*9Y* and *BRE4-uORF-out* measured in *WT* *tma20Δ*, *nmd2Δ* and *tma20Δ nmd2Δ* strains. For each allele of *BRE4* the Renilla luciferase transcript level measured in *WT* was given a value of 1 and all other measurements were corrected relative to this. RNA was isolated from cells grown in liquid culture to exponential phase at 30°C and transcript levels determined by SYBR Green RT-PCR. RNA concentrations of the samples were normalized to the loading control *BUD6*. Each measurement was performed in triplicate. Each point on the plot represents an independent measurements (mean of triplicates) and the bar represents the mean of all independent measurements.

#### 4.9 All three uORFs of *ALD3* are required to promote translation.

uORFs are generally thought to inhibit gene expression therefore we were surprised to observe that uORFs seemingly promote expression in 2 of the 22 genes we examined, *ALD3* and *CMK1*. Strikingly expression of *ALD3* was reduced by about 80% upon introduction upon abolition of the uORFs (Figure 26M). *ALD3* contains 3 uORFs that overlap each other and are, from the most proximal to *ALD3* main coding sequence 66, 30 and 75 nucleoids in length (Figure 31a).

To gain insight into the mechanism by which one or more uORF/s are seemingly required for optimal translation of *ALD3*, point mutations were individually introduced into each of the 3 uORF initiation codons, thus creating 3 new alleles of *ALD3*, termed *ALD3 – uORF1 M1V*, *ALD3 – uORF2 M1V* and *ALD3 – uORF3 M1V* (Figure 31b). *ALD3 – uORF1 M1V*, *ALD3 – uORF2 M1V*, *ALD3 – uORF3 M1V* contain point mutations in the initiation codons of the 5' most uORF, the middle uORF and the uORF closest to *ALD3* main coding sequences respectively whereas *ALD3 uORF1, 2 & 3 M1V* contains 3 point mutations in the initiation codons of the 3 uORFs.

Consistent with our previous data expression is reduced by about 80% when point mutations were introduced into all 3 uORFs in *ALD3 uORF1, 2 & 3 M1V* (Figure 31c). However introduction of a single point mutation into only the initiation codon of the 5' most uORF (*ALD3-uORF1 M1V*) resulted in only a slight reduction of *ALD3* expression and *ALD3* expression is reduced by about 40% in *ALD3-uORF2 M1V* and *ALD3-uORF3 M1V* (Figure 31c). This indicates that uORF1 has minimal effect on the expression of *ALD3* and although uORF2 and uORF3 both promote expression of *ALD3* their effect is greatly reduced compared with that of *ALD3 uORF1, 2 & 3 M1V* (Figure 31C). We investigated the regulation of *ALD3* by *TMA20* and *TMA22* however *tma20Δ* and *tma22Δ* had no significant effect on expression of any of the alleles of *ALD3* suggesting that they do not play a role in regulating *ALD3* expression (Figure 31d).

One possible mechanism by which expression of a transcript may be stimulated by a uORF is if the uORF produced a functional peptide, which enhanced expression. However if a single uORF at the *ALD3* locus encoded a functional peptide then a large decrease in *ALD3* expression would have been observed upon introduction of a point mutation a single uORF initiation codon. Another possible explanation of how *ALD3* uORFs may promote expression of a downstream ORF is if initiation at that

uORF prevented start site selection of another aberrant initiation codon, thus initiating translation of an incorrect polypeptide. This model was excluded however since there are no other uAUGs within *ALD3* TL. Translation of *ALD3* may rely on TL secondary structure or an IRES and it is possible that introducing point mutations into the TL disrupts secondary structure, thereby reducing *ALD3* expression.

At the *ALD3* locus the 5' most uORF, uORF1 is in a much stronger initiation context than uORF2 or uORF3 but *ALD3 uORF1M1V* has little impact on *ALD3* expression compared with *ALD3 uORF2 M1V* and *ALD3 uORF3 M1V*, consistent with our previous data which also indicate no correlation between a uORF initiation context and the impact that a uORF has on downstream expression.



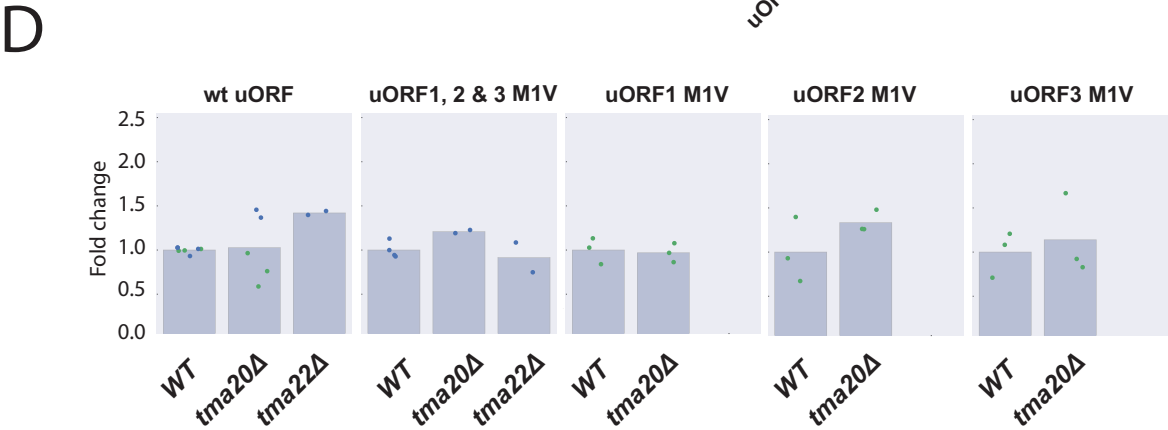
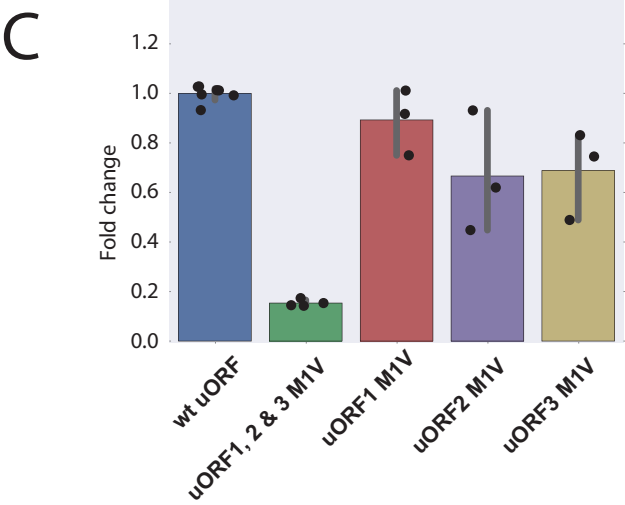
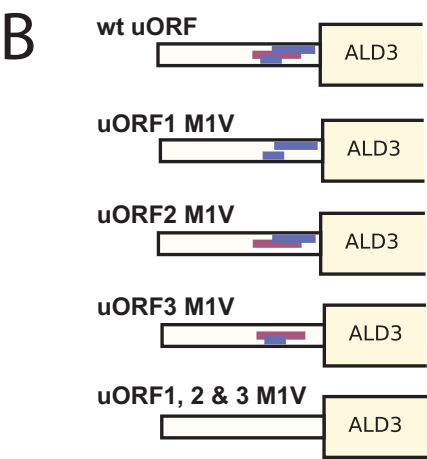
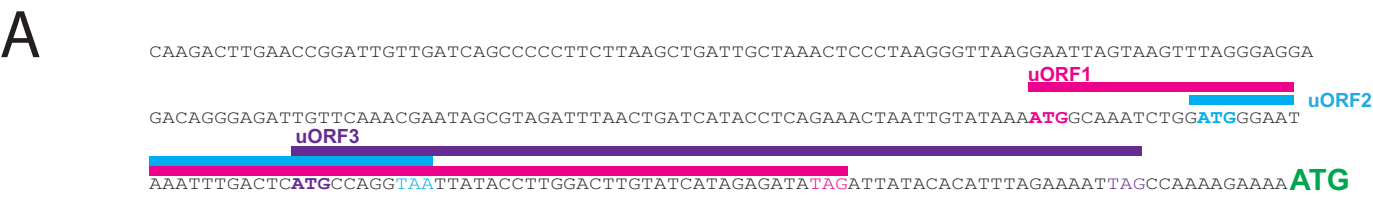


Figure 31. All three uORFs of *ALD3* are required to promote translation. A) The TL sequence of *ALD3*, indicating the positions of each of the uORFs. B) Alleles of *ALD3* used in C and D showing the uORFs that have been disrupted by the introduction of a point mutation to the initiation codon. C) The ratio of Renilla (driven by an allele of *ALD3*) over Firefly (*PGK1* driven, normalisation control) luciferase activity is calculated to obtain the normalised Renilla luciferase activity (rluc/fluc). The normalised Renilla luciferase activity that was expressed from *ALD3* (containing no point mutations) was given a value of 1 and the normalised Renilla luciferase activity that was expressed from each of the other alleles of *ALD3* was expressed relative to this. Plasmids were transformed into WT (DLY3001) cells. Each point represents a measurement from an independent colony. C) The effect, relative to *WT* of *tma20Δ*, *tma22Δ* expression of each of the alleles of each allele of *ALD3*. For each allele of *ALD3* the Renilla luciferase activity measured in *WT* was given a value of 1 and Renilla luciferase activity measured from each of the deletion strains was corrected relative to this. Cells were grown and luciferase assays conducted as per the conditions that are outlined in the 'Luciferase assay' section of the methods.

## 4.10 Discussion

Data in this chapter suggests that Tma20 and Tma22, in contrast with the reported function of their homologs in *Drosophila* (Schleich et al., 2014b), do not promote translation re-initiation in yeast. We propose an alternative model to explain how Tma20 and Tma22 can regulate expression of transcripts that contain uORFs, by reducing transcript levels. Using alleles of *BRE4* we show that Tma20 and Tma22 decrease rather than increase expression downstream of a uORF and do so by reducing transcript levels in a mechanism that is independent of NMD. We also identified, using an allele of *BRE4* that transcripts containing oORFs are also among Tma20 and Tma22 substrates.

### 4.10.1 Regulation of gene expression by uORFs

Computational analysis of the TL sequences showed that uORFs are less common than expected from random sequences and associate with reduced translation efficiencies, in agreement with published observations (Arribere and Gilbert, 2013). We selected 22 genes that contain uORFs, to test the hypothesis that Tma20 and Tma22 promote translation re-initiation. Interestingly, of these 22 genes there is substantial variation in the influence that uORFs have on gene expression. 13 genes have uORFs that repress expression, 7 have uORFs that do not affect expression and 2 have uORFs that seemingly promote expression (Figure 26). Surprisingly we found no evidence to suggest that uORFs in strong initiation contexts repress translation more than those in weaker initiation contexts (Figure 26), as has been previously predicted (Wethmar, 2014) (Figure 27e). Consistent with our observation that there is more selection against uORFs that reside in close proximity to the main coding sequence (Figure 23b), the genes whose expression was repressed the most by uORFs also contain uORFs in close proximity to the main coding sequence (Figure 27b). uORFs serve as important regulatory elements within the TL and many polymorphisms that create or eliminate uORFs have already been associated with disease (Barbosa et al., 2013a). uORFs are also more prevalent in oncogenes suggesting that understanding how they influence expression is of importance in cancer biology (Barbosa et al., 2013a). However much current understanding about how uORFs regulate expression is largely made up of gene specific mechanisms

such as *GCN4* and *CPA1*. This chapter highlights the importance of uORF-mediated control which should be considered when exploring how the expression of a gene is controlled.

#### *4.10.2 No evidence that TMA20 and TMA22 promote translation re-initiation*

The main aim of this chapter was to test the hypothesis that Tma20 and Tma22 promote translation re-initiation after translation of a uORF, as in *Drosophila* (Schleich et al., 2014b). Interestingly we found that uORF mediated regulation of expression is, in all but one case, unaffected by deletion of *TMA20* and *TMA22* (Figure 28). This suggests therefore that the model proposed in *Drosophila*, that Tma20 and Tma22 are required for optimal expression of genes that contain uORFs does not apply to yeast (Schleich et al., 2014b). More recently it has been published that mammalian homologs of *TMA20* and *TMA22* also promote expression of genes that contain uORFs but, in contrast with *Drosophila* homologs, they only promote expression of genes that have 6 nucleotide long uORFs (Schleich et al., 2017). We considered this possibility but found that Tma20 and Tma22 are not required for expression of *PGS1*, a gene that contains a 6bp uORF (Figure 28 G).

#### *4.10.3 Interaction of Tma20 and Tma22 with the ribosome has multiple consequences for translation*

Based on our analysis of alleles of *BRE4*, showing that Tma20 and Tma22 decrease expression by reducing transcript levels, we propose an alternative model to explain how Tma20 and Tma22 can affect expression of transcripts that contain uORFs. We suggest that Tma20 and Tma22 can also decrease expression by reducing transcript levels via interaction with the uORF termination codon. We propose that the interaction of Tma20 and Tma22 with a uORF can have multiple consequences for expression of the downstream ORF (Figure 32). This model is supported by biochemical evidence that MCT-1/DENR can adopt multiple roles in translation (Figure 32 a, b and c) (Skabkin et al., 2010a). MCT-1/DENR promote the dissociation of 40S from de-acetylated tRNA and mRNA after translation termination (Figure 32b) and in the same study they were also shown to promote recruitment of an acetylated tRNA to the 43S initiation complex (Figure 32c) (Skabkin et al., 2010a). The authors propose that different roles that can be adopted by Tma20/Tma22 are a

consequence of different outcomes of similar interaction of Tma20/Tma22 with the 40S and depend on the 40S ribosomal complexes (Skabkin et al., 2010a).

Figure 32 depicts the different possible outcomes that could potentially occur upon loss of Tma20 or Tma22. Published observations in mammalian extracts predict that loss of Tma20 or Tma22 would result in decreased dissociation of the 40S ribosomal complex from de-acetylated tRNA and mRNA (Skabkin et al., 2010a). This decrease in dissociation of 40S ribosomal complex from de-acetylated tRNA and mRNA after translation of a uORF could have 3 potential consequences to expression of the downstream ORF. The first is that it may increase expression by increasing translation re-initiation through maintaining the interaction of 40S with the TL (Figure 32D). The second, it may also increase expression by increasing the stability of the transcript, preventing its degradation (Figure 32E). It could also similarly decrease stability of the transcript. The third is that it may decrease expression by acting as a 'roadblock', impeding the path of other 40S complexes that otherwise would have initiated translation at that downstream ORF (Figure 32F). Published observations also suggest that loss of Tma20 or Tma22 would result in decreased recruitment of acetylated tRNA to the 43S initiation complex (Skabkin et al., 2010a), which would decrease the likelihood of translation re-initiation thereby decreasing expression (Figure 32G).

The model presented in Figure 32 suggests that the effect Tma20 or Tma22 on expression of a gene that contains a uORF is highly context dependant. This context dependence may explain why, in most of the genes that we tested, Tma20 and Tma22 had no effect on uORF-mediated regulation (Figure 28). It also offers an explanation as to the different roles of Tma20 and Tma22 homologs in mammalian cells with *Drosophila* (Schleich et al., 2014b, Schleich et al., 2017). Another possibility to explain why Tma20 and Tma22 do not promote expression of most of the genes that contain uORFs could be that they only do so under conditions of stress. Further there may be species-specific variations in their function.

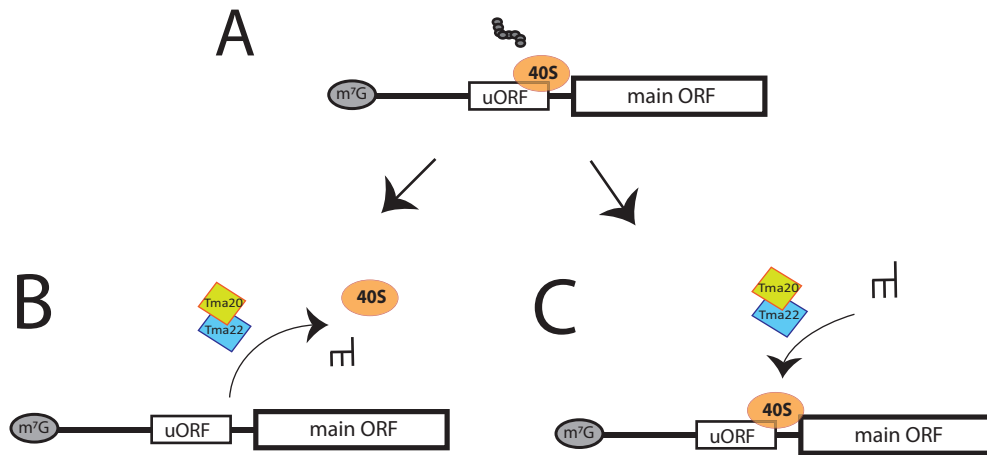
One outstanding question that remains however is why the mammalian homolog of Tma64 (eIF2D), that is proposed to also promote dissociation of 40S ribosomal complex from de-acetylated tRNA and mRNA (Skabkin et al., 2010a) has no effect on the expression of any of the genes that we tested. One explanation for this is the subtle difference in the action of Tma64 (eIF2D) and MCT-1/DENR which is that

eIF2D was shown to more strongly promote release of 40S ribosomal complex from de-acetylated tRNA and mRNA (Skabkin et al., 2010a), which possibly has different implications for expression of the downstream gene.

Our data suggests that Nmd2 decreases transcript levels of *BRE4* alleles by interaction with the uORF termination codons (Figure 30 b+c), consistent other observations that transcripts that contain uORFs are NMD substrates (Gaba et al., 2005a). Surprisingly we observed that Tma20 and Tma22 function independently of NMD to decrease transcript levels of *BRE4* (Figure 30 b+c). This implies that Tma20 and Tma22 may directly activate decapping factors or nucleases that degrade transcripts or function in a different mRNA decay pathway.

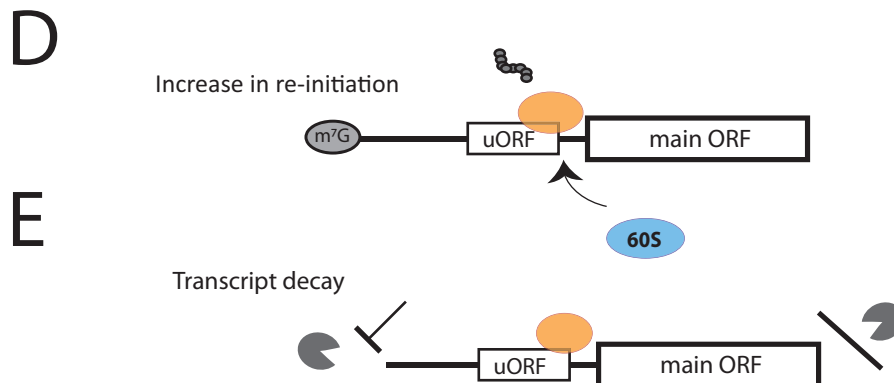
Interestingly we show that Tma20 and Tma22 also decrease the transcript levels of a *BRE4* allele that has an overlapping ORF (oORF) suggesting that *TMA20* and *TMA22* may also reduce expression of other genes that contain oORFs. The possibility that mammalian homologs of *TMA20* and *TMA22* affect the expression of genes that contain oORFs has not, that we are aware of been investigated.

# Proposed biochemical roles of Tma20/Tma22



## Possible consequences of loss of Tma20/Tma22 function

### Increase expression



### Decrease expression

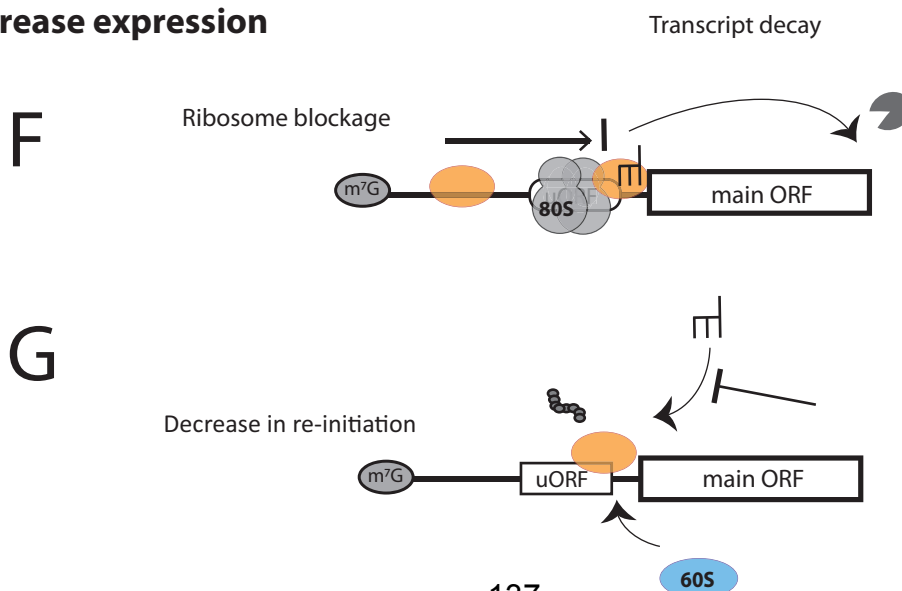


Figure 32. Proposed interactions between homologs of Tma20 and Tma22 with the post-termination ribosome. A-C) Proposed biochemical roles of mammalian homologs of Tma20/Tma22. A) following translation of a uORF and dissociation of the 60S ribosomal subunit Tma20/Tma22 may B) promote dissociation of the 40S, de-acetylated tRNA and mRNA or C) promote recruitment of an acetylated tRNA to the 40S complex. D-G) Possible consequences to gene expression of loss of Tma20 or Tma22. Following termination of translation at a uORF the 40S remains attached to the mRNA, which could result in an increase of expression of the downstream ORF by D) increased translation re-initiation, E) Decreased decay or a decrease in expression of the downstream ORF by F) resulting in a roadblock that prevent other ribosomes from reaching the main ORF or G) decreasing translation re-initiation by decreasing the recruitment of Met-tRNA<sub>i</sub><sup>Met</sup>.



## Chapter 5. Expression of *Stn1* is tightly controlled by an oORF

### 5.1 *STN1* upstream open reading frames decrease reporter gene expression

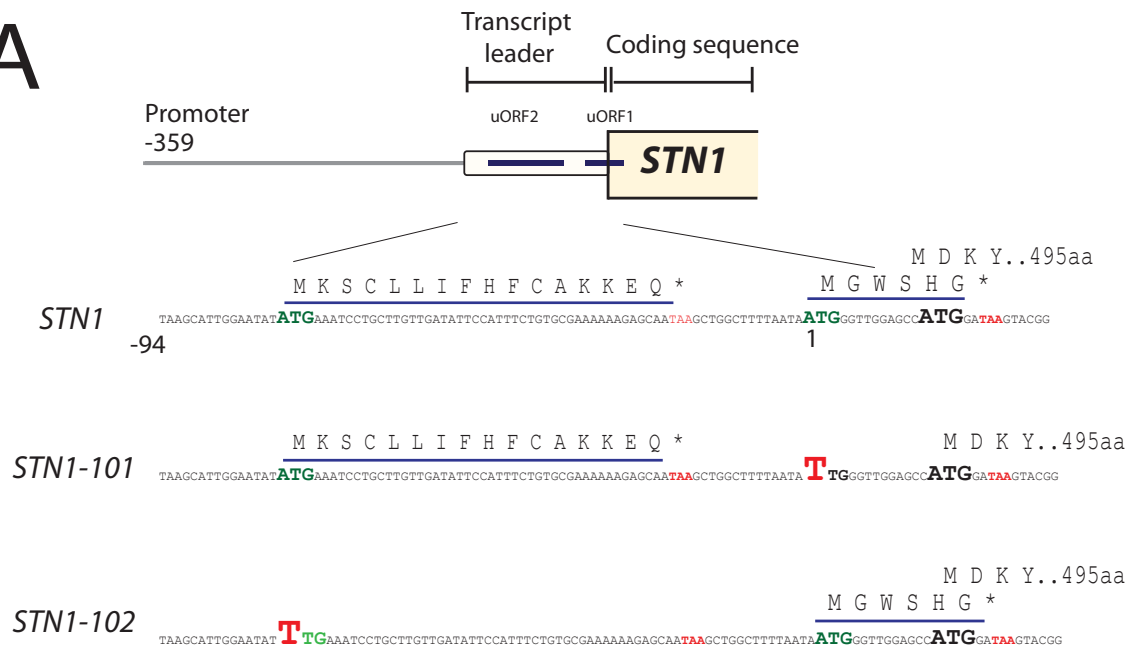
*STN1* contains two uORFs, one of which overlaps with the main ORF. There is a 16 amino acid long uORF that is positioned 27 base pairs from the main ORF start codon and a 6 amino acid long overlapping ORF that terminates 2 base pairs after the main ORF start codon (Figure 33a).

To investigate the impact of these uORFs on *STN1* expression we created two new alleles of *STN1*, each containing a point mutation in one of the uORF initiation codons and called these alleles *STN1-101* and *STN1-102* (Figure 33a). *STN1-101* contains a point mutation in the uORF closest to the main ORF and *STN1-102* contains a point mutation in the uORF most distal to the main ORF. To measure the effect each of the mutations has on expression we cloned each of the *STN1* alleles into our luciferase reporter plasmid. We found that compared with *STN1*, expression is 1.2 fold higher in *STN1-102* and remarkably, 10 fold higher in *STN1-101* demonstrating that the uORF most distal to *STN1* slightly represses expression while the overlapping uORF dramatically represses expression (Figure 33b).

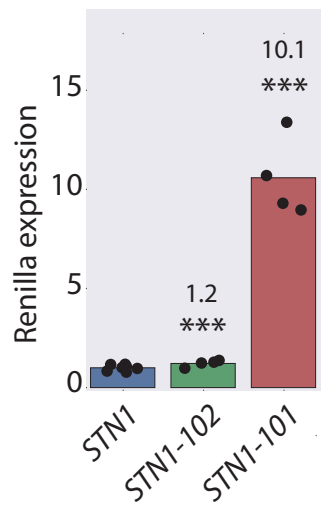
The idea that the uORF further from the main coding sequence is a much weaker suppressor of *STN1* expression than the oORF that is in closer proximity to the main coding sequence is consistent with the idea that oORFs in close proximity to the main coding sequence repress translation more than those further from the main coding sequence, as demonstrated in Figure 23.

Previous studies showed that strains ectopically overexpressing *STN1* have shorter telomeres (Dahlseid et al., 2003) and have an impaired S phase checkpoint (Gasparyan et al., 2009) suggesting that tight regulation of the *Stn1* levels are important for cell fitness. It would be of interest therefore to determine if *STN1* uORFs and oORFs regulate its expression in the genome.

# A



# B



**Figure 33 - *STN1* upstream open reading frames decrease reporter gene expression**

Figure 33 - *STN1* upstream open reading frames decrease reporter gene expression.

A) The TL sequence of *STN1*, *STN1-102* and *STN1-101* indicating the positions of the uORF and oORFs. B) The ratio of Renilla (driven by an allele of *STN1*) over Firefly (*PGK1* driven, normalisation control) luciferase activity is calculated to obtain the normalised Renilla luciferase activity (rluc/fluc). The normalised Renilla luciferase activity that was expressed from *STN1* (containing no point mutations) was given a value of 1 and the normalised Renilla luciferase activity that was expressed *STN1-102* and *STN1-101* were corrected relative to this. Plasmids were transformed into WT (DLY3001) cells. Each point represents a measurement from an independent colony. Cells were grown and luciferase assays conducted as per the conditions that are outlined in the 'Luciferase assay' section of the methods. *P* values were calculated using t-test (\*)  $P < 0.05$ , (\*\*)  $P < 0.01$ , (\*\*\*)  $P < 0.001$ .

## 5.2 *STN1* mutations strongly increase *STN1* expression from the native locus.

To confirm that Stn1 expression is negatively regulated by two uORFs we measured the levels of endogenous Stn1 in *STN1-102* and *STN1-101*. *STN1-102* and *STN1-101* were incorporated into the genome at the *STN1* locus and Stn1 was fused with a Myc epitope tag to allow for detection by western blot.

In agreement with our data from the luciferase assay, Stn1 levels are about 10 fold higher in *STN1-102* compared with *STN1* demonstrating that *STN1* expression is dramatically attenuated by the oORF (Figure 34a +b). This increase in Stn1 is accompanied by a much smaller increase, of only 4 fold, in mRNA level demonstrating that the oORF decreases the level of *STN1* transcript in addition to reducing translation (Figure 34c).

Stn1 levels in *STN1-101* are similar to wild type *STN1*, which, in contrast with the luciferase assay data, suggests that the distal uORF does not repress *STN1* expression at least to a level detectable by western blot (Figure 34a +b). There was a small increase in the level of *STN1-101* mRNA, however this was not found to be statistically significant (Figure 34c).

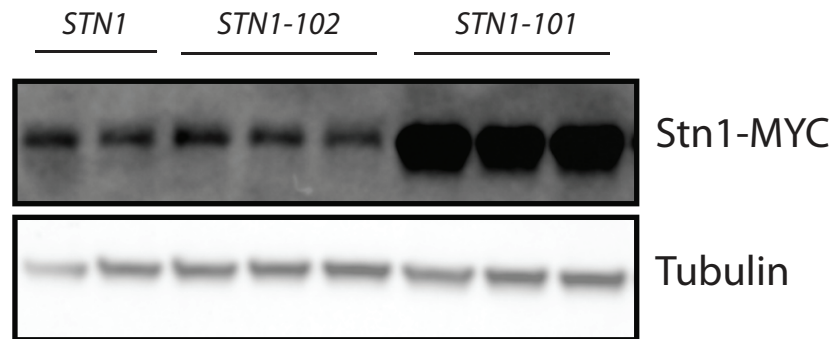
It is unclear why there was an increase in *STN1-102* expression observed by luciferase assay but not by western blot. One possible explanation could be due to the sensitivities of the luciferase assay compared to western blot in that the 1.2 fold increase in *STN1-102* expression detected by luciferase assay is below the detection limit of the western blot. Another possible explanation is that the increase in *STN1-102* expression detected by luciferase assay is an artefact of the luciferase assay.

The observation that the oORF decreases *STN1* transcript levels suggests that the oORF may be reducing the stability of the transcript, raising the question of what mechanism may degrade mRNAs that have oORFs. One possibility is the NMD pathway. A recent study found NMD substrates to have significantly lower ribosome densities and higher rates of out of frame translation (Celik et al., 2017a). *STN1* transcript is presumably associated with less ribosomes since it produces substantially less Stn1 compared with *STN1-101*, therefore may be targeted by NMD more than *STN1-101*. Additionally out of frame translation will be dramatically reduced on *STN1-101*, which does not have an oORF, thereby making *STN1* again more likely to be targeted by NMD.

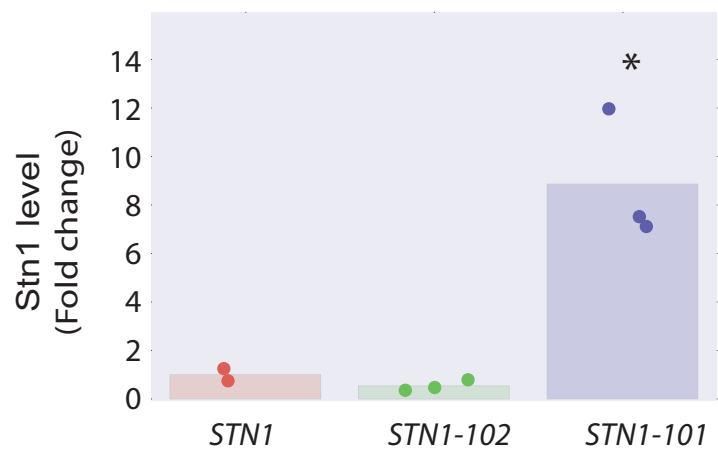
The dramatic increase in Stn1 levels observed in *STN1-101* indicates the oORF plays an important regulatory role in maintaining appropriate levels of Stn1.

In the course of our studies we noticed that *STN1-101-MYC* strains contains an additional point mutation within the main coding sequence of *STN1*. The increase in Stn1 levels observed in *STN1-101-MYC* are similar to those observed in our luciferase, suggesting that they are unlikely a consequence of the secondary mutation. However measurements of Stn1 and *STN1* transcript levels in *STN1-101-MYC* should be repeated.

A



B



C

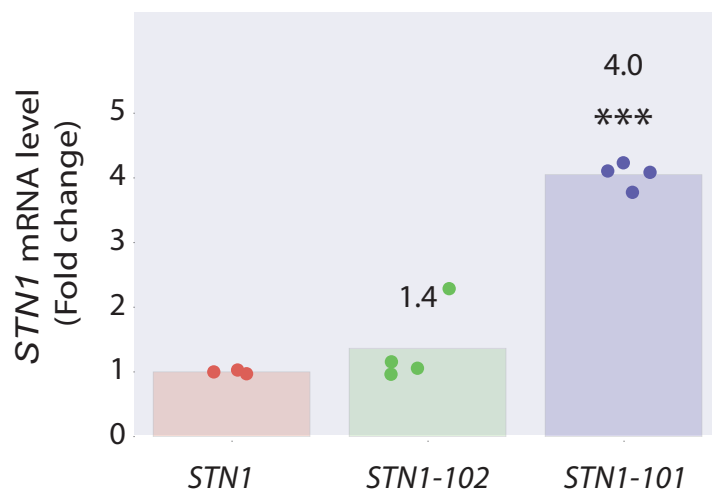


Figure 34 - *STN1* mutations strongly increase *STN1* expression from the native locus. A) Western blot analysis of Stn1 protein levels in *STN1*, *STN1-102* and *STN1-101*. Strains contain Stn1-13Myc. Proteins were extracted from cells that were grown in liquid culture to exponential phase at 30°C. Antibodies against Myc-tag were used to detect Stn1 and antibodies against tubulin were used for loading control. B) Quantification of the western blot. C) Transcript levels of Stn1-13Myc measured in *STN1*, *STN1-102* and *STN1-101*. RNA was isolated from cells grown in liquid culture to exponential phase at 30°C and transcript levels determined by SYBR Green RT-PCR. RNA concentrations of the samples were normalized to the loading control *BUD6*. The mean of 3 *STN1* strains was given the value of 1 and all other values were corrected relative to this. Each measurement was performed in triplicate. Each point on the plot represents an independent measurements (mean of triplicates) and the bar represents the mean of the 4 independent measurements. *P* values were calculated using t-test (\*)  $P < 0.05$ , (\*\*)  $P < 0.01$ , (\*\*\*)  $P < 0.001$

### 5.3 *STN1-101* and *STN1-102* improve fitness of cells with *cdc13-1* telomere defects.

Stn1, together with Ten1 and Cdc13 forms the CST complex that functions in telomere capping and telomerase regulation (Sun et al., 2009). Overexpression of *STN1* improves fitness of *cdc13-1*, presumable because Stn1 can compensate for loss of telomere capping in *cdc13-1* (Grandin et al., 1997). We therefore wondered if the uORF and particularly the oORF of *STN1* might contribute to the *cdc13-1* fitness defect.

*STN1-102* and *STN1-101* were combined with *cdc13-1* and fitness of resulting strains was assessed by spot test. We controlled our experiments with strains that contain a *URA3* at the *STN1* locus as in *STN1-101* and *STN1-102* to account for any possible side effects of the marker. Thus *STN1*, *STN1-101* and *STN1-102* differ by single point mutations. The growth of *cdc13-1*, at 26°C, is marginally improved in strains that contain a *URA3* at the *STN1* locus suggesting that the marker may have a minor affects on levels of Stn1 (Figure 35). Interestingly however, *cdc13-1 STN1-102* grows better than *cdc13-1* and importantly *cdc13-1 STN1* at 26°C suggesting the uORF closest to the 5' cap may slightly reduce *STN1* levels, supporting the data from our luciferase assay (Figure 35). Remarkably, growth of *cdc13-1* is dramatically improved by *STN1-101* at 26°C-29°C, in line with our observation that *STN1-101* results in a substantial increase in the level of Stn1 (Figure 35). Our results are consistent with published data showing that overexpression of *STN1* is able to partially suppress *cdc13-1* (Grandin et al., 1997).

Stn1 has been shown to completely compensate for loss of Cdc13 when delivered to the telomere by being fused to the DNA binding domain of Cdc13 (Pennock et al., 2001). However *STN1-101* does not improve fitness of *cdc13-1* at temperatures above 29°C indicating that there is still a requirement for Cdc13, possibly in recruitment of Stn1 to the telomeres.



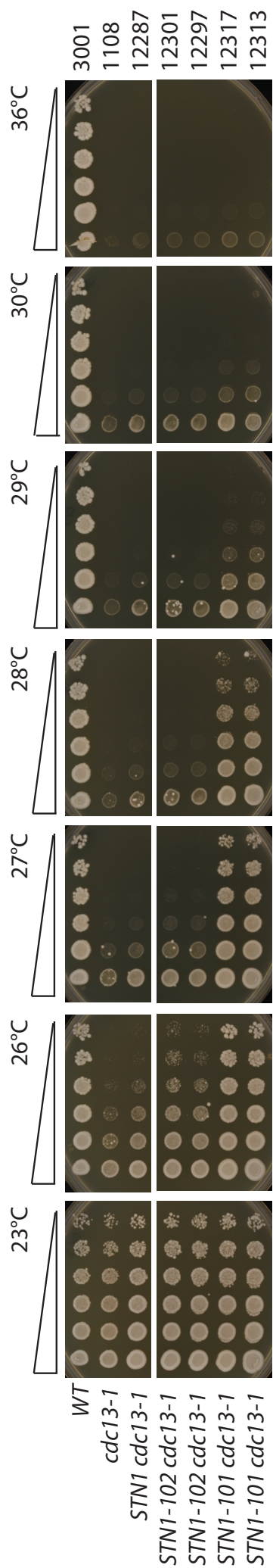


Figure 35 - *STN1-101* and *STN1-102* improve fitness of cells with *cdc13-1* telomere defects. Yeast strains of the indicated genotypes were grown to saturation in liquid YEPD, serially diluted, 5-fold, in water, and spotted onto YPD solid media. Plates were incubated for two days at the indicated temperatures before being photographed. Genotypes and strain numbers are listed on either side of the spot tests.

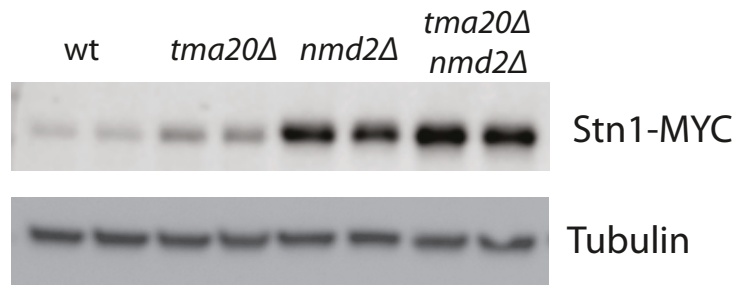
#### 5.4 *TMA20* and *NMD2* reduce *Stn1* levels

Deletion of nonsense mediated decay genes (NMD) increase fitness of *cdc13-1* and it is thought that this is partially due to the higher levels of *Stn1* in *NMD* null strains (Addinall et al., 2011, Enomoto et al., 2004). *TMA20* and *TMA22* function in the same pathway as *NMD* genes to affect fitness of *cdc13-1* leading us to hypothesise that *TMA20* and *TMA22* also affect fitness of *cdc13-1* by increasing levels of *Stn1* (Figure 15).

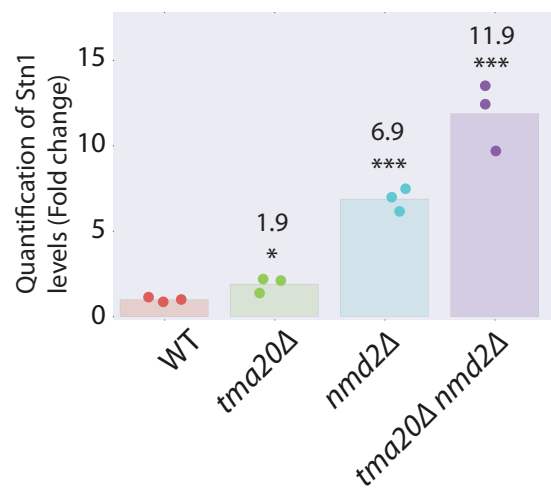
Consistent with our hypothesis levels of *Stn1* are higher in *tma20Δ*, demonstrating that *TMA20* inhibits *Stn1* levels (Figure 36a). The levels of *Stn1* in *tma20Δ* are lower than in *nmd2Δ*, in line with the observation that *nmd2Δ* is a much stronger suppressor of *cdc13-1* thermo-sensitivity than *tma20Δ*. While *tma20Δ* results in an increase in *Stn1*, there was not a statistically significant increase in *STN1* transcript suggesting that that *TMA20* decrease the translation of *Stn1* (Figure 36b). On the other hand *nmd2Δ* results in a similar increase in *STN1* mRNA and protein level suggesting that the effect of NMD on *Stn1* levels are due to NMD decreasing *Stn1* transcript levels, consistent with the role on NMD in mRNA decay (Figure 36b). The levels of *Stn1* protein and transcript that we observe in *nmd2Δ* accurately reproduce those previously published (Addinall et al., 2011). Interestingly levels of *Stn1* are slightly higher in *tma20Δ nmd2Δ* compared with *tma20Δ* or *nmd2Δ*, indicating that *TMA20* and *NMD2* act in different pathways to decrease *Stn1* (Figure 36a). However transcript levels of *tma20Δ nmd2Δ* and *nmd2Δ* are similar consistent with the idea the effect of *tma20Δ* on *Stn1* levels is due to increased translation (Figure 36b).

The NMD pathway is known to promote the degradation of transcripts that containing uORFs suggesting that NMD may repress *Stn1* levels as a consequence of the *Stn1* uORFs (Arribere and Gilbert, 2013). Observations that NMD targets poorly translated transcripts that have higher rates of out of frame translation support the idea that NMD targets *STN1* as a consequence of the oORF (Celik et al., 2017a). The oORF of *STN1*, in addition to reducing translation efficiency also reduces the *STN1* transcript levels, providing further evidence that NMD could reduce *Stn1* levels via the oORF. Some of our data also indicate that *TMA20* contributes the reduction of expression that occurs as a result of an oORF leading us to hypothesise that *TMA20* may also repress *STN1* via the oORF.

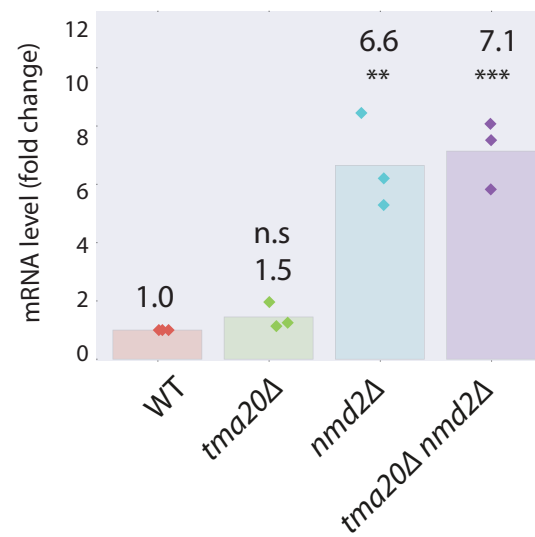
A



B



C



### Figure 36 - *TMA20* and *NMD2* reduce *Stn1* levels

Figure 36. *TMA20* and *NMD2* reduce *Stn1* levels. A) Western blot analysis of *Stn1* protein levels in *WT*, *tma20Δ*, *nmd2Δ* and *tma20Δ nmd2Δ* strains contain *Stn1*-13Myc. Proteins were extracted from cells that were grown in liquid culture to exponential phase at 30°C. antibodies against Myc-tag were used to detect *Stn1* and antibodies against tubulin were used for loading control. B) Quantification of the western blot. C) Transcript levels of *Stn1* measured in *WT*, *tma20Δ*, *nmd2Δ* and *tma20Δ nmd2Δ* strains. RNA was isolated from cells grown in liquid culture to exponential phase at 30°C and transcript levels determined by SYBR Green RT-PCR. RNA concentrations of the samples were normalized to the loading control *BUD6*. Each time the experiment was performed the *WT* was given a value of 1 and other measurements were expressed relative to this. Each measurement was performed in triplicate. Each point on the plot represents an independent measurements (mean of triplicates) and the bar represents the mean of the independent measurements. *P* values were calculated using t-test (\*) *P* < 0.05, (\*\*) *P* < 0.01, (\*\*\*) *P* < 0.001

## 5.5 *TMA20* and *NMD2* reduce *Stn1* levels via the overlapping uORF.

Deletion of *TMA20* and *NMD2* or ablation of the oORF all results in increased levels of *Stn1*. We hypothesised that *TMA20* and *NMD2* may function through the oORF to reduce levels of *Stn1* and used a luciferase assay to test this.

The effect of *tma20Δ* and *nmd2Δ* on the level of reporter gene produced from *STN1*, *STN1-101* and *STN1-102* was measured. In agreement with the western blot analysis of *Stn1*, deletion of *NMD2* and *TMA20* result in an increase in *STN1* expression, which is more profound in *nmd2Δ* (Figure 37). Similarly *tma20Δ* and *nmd2Δ* also resulted in an increase in expression of *STN1-102*. However *tma20Δ* and *nmd2Δ* have no impact on expression of *STN1-101* implying that *TMA20* and *NMD2* repress *STN1* expression via the overlapping uORF (Figure 37).

Our data support a model the reduction in *Stn1* that occurs in *NMD2* and *TMA20* null alleles is a consequence of *Stn1* oORF. NMD and the oORF both cause a reduction in the transcript levels of *STN1*, supporting the idea that the oORF results degradation of the transcript via NMD. This could be a consequence of ribosomes stalling on the oORF or a reduced density of ribosomes on the main ORF of *STN1*. However a mechanism to explain how the oORF of *STN1* causes a *Tma20* mediated reduction in expression remains to be established. Our data indicate that *TMA20* decreases expression of *Stn1* independently of NMD, and also that *TMA20* reduces the translation of *STN1*.

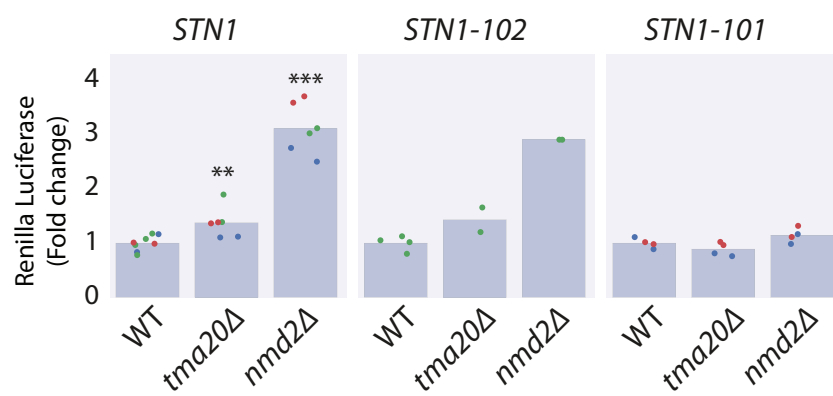


Figure 37. *TMA20* and *NMD2* reduce *Stn1* levels via the overlapping uORF. The effect of *tma20Δ* or *nmdΔ* is quantified in each allele of *STN1* (*STN1*, *STN1-102* or *STN1-101*). The ratio of Renilla (driven by an allele of *STN1*) over Firefly (*PGK1* driven, normalisation control) luciferase activity is calculated to obtain the normalised Renilla luciferase activity (rluc/fluc). For each allele of *STN1* the Renilla luciferase activity measured in WT was given a value of 1 and Renilla luciferase activity measured from *tma20Δ* or *nmdΔ* were corrected relative to this. The points on the graph represent independent measurements and the bars are a mean of all the independent measurements. Individual measurements are coloured according to the time that they were obtained, so point of the same colour were obtained on the same experiment. Cells were grown and the luciferase assay conducted according to the 'Luciferase assay' section in the methods. *P* values were calculated using t-test (\*)  $P < 0.05$ , (\*\*)  $P < 0.01$ , (\*\*\*)  $P < 0.001$



## 5.6 Genetic evidence supports the idea that *TMA20* decreases fitness of *cdc13-1* by decreasing levels of Stn1

*tma20Δ* and *nmd2Δ* both increase the levels Stn1 and increased levels of Stn1 increase the fitness of *cdc13-1*. Together this forms the hypothesis that the increase in fitness of *cdc13-1* that occurs upon deletion of *TMA20* or *NMD2* is a result of their influence on levels of Stn1. This idea has been proposed previously to explain the effect of *nmd2Δ* on *cdc13-1* fitness (Addinall et al., 2011). However the extent to which *NMD2* and *TMA20* act through *STN1* to increase fitness of *cdc13-1* remains to be tested.

To investigate the extent to which *NMD2* and *TMA20* act through *STN1* to increase fitness of *cdc13-1* we combined *STN1-101* and *STN1-102* with *tma20Δ* and *nmd2Δ* in a *cdc13-1* background and measured fitness by spot test. *tma20Δ* does not improve fitness of *STN1-102* or *STN1-101* in the context of *cdc13-1*, supporting the idea that that *tma20Δ* improves fitness of *cdc13-1* by increasing levels of Stn1 (Figure 38- compare lines 3, 4 & 5 with 6 & 7). On the other hand, *nmd2Δ* strongly improves the fitness of both *STN1-102* and *STN1-101* in a *cdc13-1* background (Figure 38- compare lines 3, 4 & 8 with 9 & 10).

The increase in fitness of *cdc13-1 STN1-101* observed upon deletion of *NMD2* could be because *nmd2Δ* might further increases the level of Stn1 produced in *cdc13-1 STN1-101*, thus resulting in an increase in fitness. However this contradicts our data from the luciferase assay that suggests that *nmd2Δ* has no effect on the level of Stn1 produced from *STN1-101*. An alternative explanation is that effect of *nmd2Δ* on growth of *cdc13-1* is only partially due to an increase in Stn1. In support of this idea, *nmd2Δ* also increases the level Ten1 (Enomoto et al., 2004) and co-overexpression of *TEN1* and *STN1* improves *cdc13-1* fitness more than the individual overexpression of either gene (Grandin et al., 2001b).

There is no observable difference in the fitness of *tma20Δ* and *STN1-102* in a *cdc13-1* background despite the higher levels of Stn1 in *tma20Δ* compared with *STN1-102*. This indicates that there is not a linear correlation between the level of Stn1 and suppression *cdc13-1* temperature sensitivity. In agreement with this experiments

from our lab have shown equal growth of *cdc13-1* cells transformed with either a low or high copy plasmid expressing *STN1*.

In summary our data suggests that *TMA20* decreases fitness of *cdc13-1* by decreasing levels of Stn1 whereas the decrease in *cdc13-1* fitness that occurs as a result of *NMD* is only partially explained by *NMD* decreasing levels of Stn1.

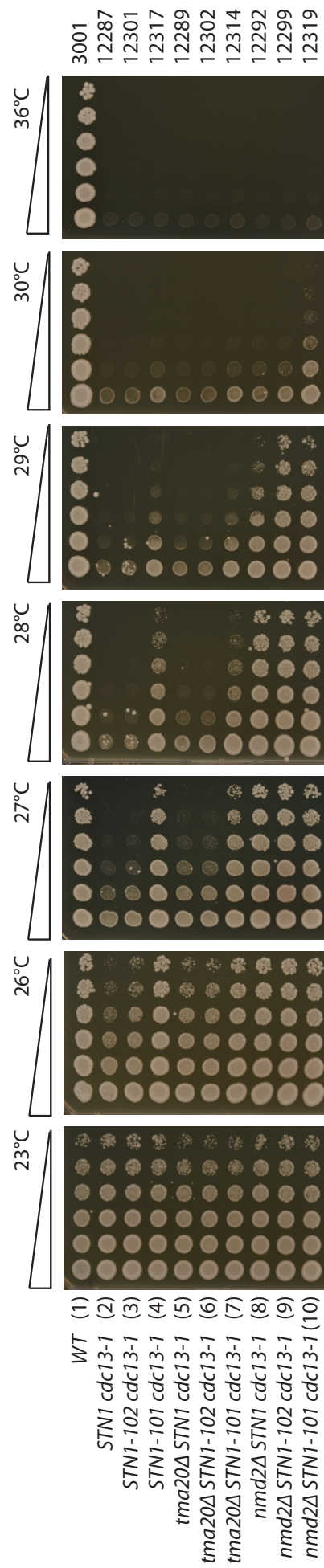


Figure 38. Genetic evidence supports the idea that *TMA20* decreases fitness of *cdc13-1* by decreasing levels of Stn1. Yeast strains of the indicated genotypes were grown to saturation in liquid YEPD, serially diluted, 5-fold, in water, and spotted onto YPD solid media. Plates were incubated for two days at the indicated temperatures before being photographed. Genotypes and strain numbers are listed on either side of the spot tests.

## 5.7 *STN1-102* and *STN1-101* do not affect sensitivity to HU

*STN1* overexpression has been shown to disrupt the S-phase checkpoint leading to HU and MMS sensitivity (Gasparyan et al., 2009). This led us to hypothesise that the role of the oORF in maintaining low levels of Stn1 may be important to maintain a functional S phase checkpoint.

The sensitivity of *STN1-102* and *STN1-101* to HU was assessed by spot test. For comparison, *rad50Δ* was included, since it is sensitive to HU (Shor et al., 2002). Surprisingly we observe no difference in the growth of *STN1-101* compared with WT or *STN1-102* when strains were spotted onto HU, despite the obvious growth defect of the HU sensitive control, *rad50Δ*.

The observation that *STN1-101* does not reduce sensitivity to HU is surprising since previous studies show that overexpression of *STN1* sensitises cells to HU (Gasparyan et al., 2009). One possible explanation for these differences might be due to differences in strain backgrounds. Alternatively it could be an artefact of the differences between overproducing Stn1 via *STN1-101* and artificially expression from a plasmid.

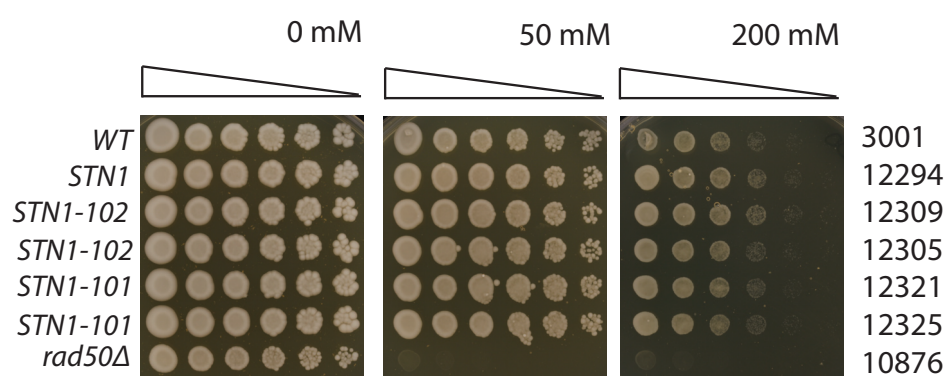


Figure 39. *STN1-102* and *STN1-101* do not affect sensitivity to HU. Yeast strains of the indicated genotypes were grown to saturation in liquid YEPD, serially diluted, 5-fold, in water, and spotted onto YPD solid media containing indicated concentration of HU. Plates were incubated for two days at 30°C before being photographed. Genotypes and strain numbers are listed on either side of the spot tests.

## 5.8 Discussion

In this chapter we describe a novel mechanism of regulating Stn1 levels whereby they are dramatically reduced by an oORF. We also show that Tma20, via this oORF, decrease levels of Stn1, thus explaining how Tma20/Tma22 decreases fitness of *cdc13-1*. We show that NMD, which has been known for some time, to regulate Stn1 expression (Dahlseid et al., 2003) (Advani et al., 2013), decreases Stn1 levels via the oORF. Our data suggests that while *TMA20* decreases the fitness of *cdc13-1* by decreasing levels of Stn1, the increase in fitness of *cdc13-1* that occurs in *nmd2Δ* is only partially explained by the increased levels of Stn1 in these strains.

### 5.8.1 *Stn1* levels are tightly controlled by an oORF

*STN1* contains two uORFs, one of which overlaps with the main oORF (Figure 33a). To enable us to study the influence of these uORFs on Stn1 levels we created two new alleles of *STN1*. *STN1-102* harbours a point mutation in the initiation codon of the uORF most distal to the main ORF and *STN1-101* harbours a point mutation in the initiation codon of the overlapping ORF (Figure 33a). While our luciferase assay indicates that *STN1-102* produces increased levels of Stn1, this result was not reproduced when we introduced *STN1-102* into the genome and measured endogenous levels of Stn1 (Figure 33b, Figure 34). Remarkably *STN1-101* dramatically increase levels of Stn1 indicating that the oORF substantially reduces *STN1* expression (Figure 34). We showed that the oORF reduces Stn1 levels by a combination of reducing transcript levels and decreasing translation and our data suggest that the decrease in transcript levels are due to NMD (Figure 34).

Translation is an extremely effective mechanism of regulating gene expression in response to environmental stimulus since it allows immediate and selective changes in protein levels. There are many examples of translational control mechanisms that are mediated by uORFs, especially mechanisms, such as the uORF that facilitate expression of *GCN4* in response to stress. In *GCN4* following translation of the first or second uORF the ribosome reinitiates translation at a downstream uORF and subsequently dissociates from the transcript (Hinnebusch, 2005). However in conditions of stress the ribosome will resume translation instead at the main coding sequence, thus up-regulating expression of Gcn4 (Hinnebusch, 2005). It would be of



interest to determine if the uORFs of *STN1* regulate expression by increasing translation under conditions of stress. It is tempting to hypothesise that a similar mechanism of translational control could occur on STN1 where following re-initiation after translation of the uORF furthest from the main ORF, initiation either occurs on the oORF, thereby decreasing *STN1* expression of the main ORF which would increase the level of *STN1*.

Interestingly the human version of *STN1* also contains an overlapping ORF that has a termination codon that overlaps with the main ORF initiation codon. The oORF in human STN1 is 15 amino acids, which is slightly longer than the 6 amino acid long oORF in yeast (Figure 40). The overlapping ORF in human STN1 contains two initiation sites, which presumably increases the likelihood of capturing a scanning ribosome. Since the human STN1 also contains an oORF the mechanism of translational regulation may be conserved.

#### 5.8.2 *Degradation of Stn1 by NMD*

It has been known for some time that NMD decreases *STN1* transcript levels but conflicting explanations have been proposed to describe the mechanism by which NMD reduced Stn1 levels (Dahlseid et al., 2003) (Advani et al., 2013). It has previously been demonstrated that the promoter region of *STN1* is sufficient to confer regulation by NMD (Dahlseid et al., 2003). Our data builds on this hypothesis by demonstrating that indeed the URS is responsible for NMD mediated regulation of *STN1* expression but shows that it is in fact the oORF rather than the promoter that mediates this regulation (Figure 37). Our data however contradicts some of the conclusions of this paper, which show that NMD doesn't affect the half-life of *STN1* transcript, thus implying that the effects of NMD on Stn1 are indirect (Dahlseid et al., 2003). Our model on the other hand implies that the effects of NMD on Stn1 levels are via a direct interaction of NMD factors with the oORF. A separate study suggests that NMD degradation of Stn1 is due to a programmed ribosomal frameshift (PRF) sequence within the coding sequence of *STN1* (Advani et al., 2013). There could therefore be more than one mechanism to explain how NMD is recruited to Stn1.

### 5.8.3 *Tma20 and Tma22 decrease fitness of cdc13-1 by decreasing Stn1 levels*

The main aim of this thesis was to determine how *tma20Δ* and *tma22Δ* increase fitness of *cdc13-1*. We propose a model based on data from this chapter that Tma20 and Tma22 affect fitness of *cdc13-1* and *stn1-13* by increasing the level of Stn1 in a mechanism that is dependant on the oORF of Stn1. Data from Chapter 3 shows that *TMA20* act in the same pathway as NMD to decrease fitness of *cdc13-1* (Figure 15). Although Tma20 and NMD both increase fitness of *cdc13-1* by increasing levels of Stn1 our data also suggest that Tma20 and NMD affect oORF-mediated repression of Stn1 by different mechanisms (Figure 37). NMD decreases *STN1* transcript levels while Tma20 decreases the translation of *STN1*. This is in agreement with some of our observations from Chapter 4 showing that NMD and Tma20/Tma22 independently decrease oORF-mediated expression of *BRE-uORF-out* (Figure 30). The mechanism by which Tma20 reduces the translation of Stn1 is as yet unclear, although we show that it is dependant on the oORF. One possible mechanism could be by increasing the translation re-initiation following translation of the oORF. Since the oORF termination is only 2 nucleotides downstream of the main initiation codon of STN1 it may be possible that the ribosome migrates across these 2 nucleotides to reach the main initiation codon.

### 5.8.4 *Tma20 and Tma22 regulate expression of genes with oORFs*

We have now shown that Tma20 regulates expression of two genes, *STN1* and *BRE4*, by a mechanism that is dependent on an oORF. However regulation of Stn1 by Tma20 occurs by a decrease in translation while regulation of Bre4 by Tma20 occurs as a result of a decrease in transcript levels. This is compatible with our model proposed in chapter Chapter 4 (see Discussion 4.10.3 and Figure 32) that predicts that the interaction of Tma20/Tma22 with the ribosome following translation termination on a uORF or oORF has multiple possible consequences to gene expression. It would be of interest to find out if Tma20/Tma22 regulate the expression of other genes that contain oORFs and further to test if in mammalian cells they regulate expression of this group of genes.

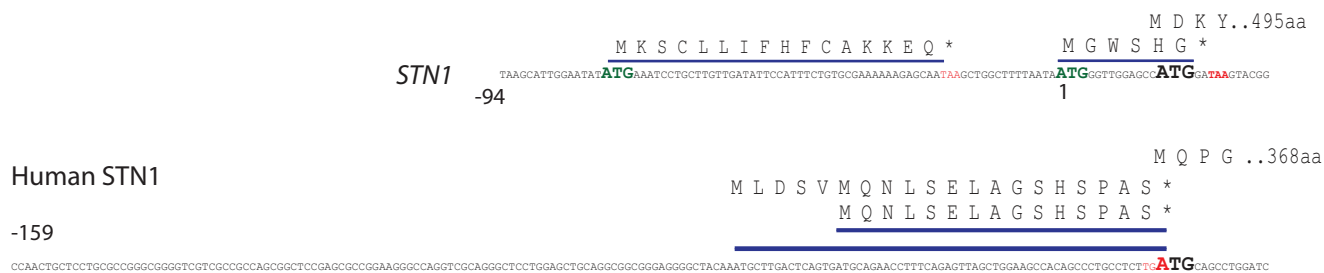


Figure 40: Comparison of the yeast and human *STN1* TL sequences. The ORFs are indicated.

## Chapter 6. General Discussion

### 6.1 Tma20 and Tma22 decrease fitness of *cdc13-1* by reducing levels of Stn1

The main aim of this thesis was to understand how two interacting proteins, Tma20 and Tma22, affect the fitness of cells with *cdc13-1* telomere capping defect. We answered this question, demonstrating that Tma20 and Tma22 decrease fitness of *cdc13-1* by reducing levels of the telomere capping protein Stn1. Further we provide a mechanism to explain how Tma20 and Tma22 reduce the level of Stn1, which is by increasing the repressive effect of the oORF. Remarkably, *tma20Δ* has no effect on the fitness of *cdc13-1* when a single base pair mutation is introduced into the initiation codon of an oORF at the *STN1* locus (*STN1-101*) (Figure 38). This suggests that Tma20 decreases fitness of *cdc13-1* via interaction with this oORF. Consistent with this we demonstrate that Tma20 only reduces expression of *STN1* when expressed from WT *STN1* allele that contains the oORF, but not from *STN1-101* (Figure 37). The human homolog of Tma20 de-regulates cell cycle checkpoints and increases genome instability (Hsu et al., 2007). Therefore it is possible that in humans MCT-1 may result in increased levels of Stn1, which contributes to genome instability and possibly other oncogenic phenotypes that are induced by MCT-1.

### 6.2 Genetic interactions between *TMA20* and *NMD2*

Overexpression of Stn1 or Ten1 and the deletion of NMD factors have all been previously shown to increase the fitness of *cdc13-1* (Enomoto et al., 2004). This, in combination with the observation that *nmd2Δ* cells have elevated levels of Stn1 and Ten1 form the hypothesis that *nmd2Δ* improves fitness of *cdc13-1* at least in part by increasing levels of Stn1 and Ten1 (Dahlseid et al., 2003) (Enomoto et al., 2004). We build on this knowledge showing that the oORF of *STN1* is what causes it to be an NMD substrate. We are also able to show that the increase in expression of *cdc13-1* is indeed only partially explained by an increase in the level of Stn1 since deletion of *NMD2* improves fitness of *cdc13-1 nmd2Δ* (Figure 38). It would be interesting to test if overexpression of *TEN1* in *STN1-101* phenocopies the *cdc13-1* suppression that is observed in *STN1-101 nmd2Δ*. This would address the question of whether the increase in fitness of *cdc13-1* observed when the NMD pathway is perturbed is a

solely the result of only increased Stn1 and Ten1 levels or if NMD also affects another pathway that contributes to fitness of *cdc13-1*. Further *STN1-101* results in higher Stn1 levels than *nmd2Δ* suggesting that overexpression of *TEN1* in *STN1-101* may even permit growth of *cdc13Δ*. *cdc13Δ* strains are normally inviable but combinations of certain gene deletions have been shown to rescue *cdc13Δ* (Holstein et al., 2014).

One seemingly contrasting observation is that *NMD2* and *TMA20* act in the same pathway to increase the fitness of *cdc13-1* (Figure 15) and we conclude that this is by decreasing levels of Stn1 (Figure 36), but *NMD2* and *TMA20* decrease the levels of Stn1 by different mechanisms (Figure 36). We reason that this because increase in the level of Stn1 that occurs in *nmd2Δ tma20Δ* compared with *nmd2Δ* is not large enough to further increase the fitness of *cdc13-1*.

### **6.3 Levels of Stn1 are controlled by an oORF**

We show that levels of Stn1 are tightly regulated by an oORF and, to a much lesser extent a uORF (Figure 34). The oORF dramatically reduces levels of Stn1 by reducing translation and decreasing the levels of transcript (Figure 34). We demonstrate biological relevance for the oORF, which dramatically decreases fitness of *cdc13-1* (Figure 35). The question still remains however of what advantage this oORF confers to cells to have allowed it to be selected for. Most oORFs are heavily selected against, as we demonstrated in (Figure 21). Interestingly overexpression of *STN1* has previously been shown to result in decreased telomere length (Dahlseid et al., 2003) suggesting that maintaining low levels of Stn1 are important for appropriate telomere length regulation.

There are many examples of uORFs functioning as regulatory elements that facilitate the expression of genes under conditions of stress, some of which are discussed in Chapter 1: Section 1.9.3. It possible therefore the oORF and/or uORF of *STN1* facilitate the expression of *STN1* in response to of stress. We, and others have demonstrated that increased levels of Stn1 improve the fitness of cells with *cdc13-1* telomere defect but it is not know if Stn1 levels increase as a response to *cdc13-1*. It would be interesting to explore if Stn1 levels increase in response to any environmental or endogenous stresses (such as a gene mutation). Another approach

would be to test if *STN1-101* affects the fitness of cells in response to stresses other than *cdc13-1* or HU.

Control over telomere protein levels is critically important to cancer (Jafri et al., 2016). Most cancer cells aberrantly express telomerase, which allows them to divide indefinitely (Iyer et al., 2005) and consistent with this point mutations in the telomerase subunit hTERT promoter are among the most frequent mutations in human carcinogenesis (Bell et al., 2016). Cancer cells that do not express telomerase continue to proliferate by using telomerase independent mechanisms to maintain telomere length, termed Alternative lengthening of Telomeres (ALT) (Gocha et al., 2013). Interestingly Stn1 has been implicated in facilitating the switch to ALT since in *K. lactis* a mutation in *STN1* has been shown to produce an ALT-like telomere phenotype. Further mutations in Stn1 are associated with Coats plus syndrome, a disease that is characterised by short telomeres (Simon et al., 2016). There is also evidence to suggest that Stn1 is important for cancer cells since as suppression of STN1 increases the sensitivity of cancer cells to a range of genotoxins used in chemotherapy (Zhou and Chai, 2016). It seems likely that the tight regulation of STN1 levels in humans is also important and since STN1 in humans also has an oORF (Figure 40) suggesting that the mechanism of regulation may be conserved. Since *STN1-101* induces the overexpression of Stn1 we predict that this mutation is dominant however it would be of interest to test this. Dominant mutations in human disease are more harmful since only one copy is required to confer the disease phenotype. An important future experiment is to confirm that this mutation is dominant in yeast.

#### **6.4 Multiple ways that Tma20 and Tma22 can affect expression of genes that contain uORFs or oORFs**

An aim of this thesis was to learn about the function of Tma20 and Tma22 of which there is little understanding of. Since they were discovered as factors that interact with the 40S ribosomal subunit, there has been nothing published about their function. We tested therefore, the hypothesis that they promote translation re-initiation, as was shown in *Drosophila* (Schleich et al., 2014b). We tested this hypothesis using 22 genes that contain uORFs (Figure 25). However we found no evidence to suggest that they affect translation in a manner dependant on the uORF/s in 21 of these cases (Figure 28). Further we found evidence to suggest that

the mechanism by which Tma20 and Tma22 mediated expression of the remaining gene *BRE4*, is not by translation re-initiation but rather by decreasing transcript levels (Figure 30). Interestingly the possibility that in *Drosophila* Tma20 and Tma22 affect transcript levels rather than promote translation re-initiation was not tested (Schleich et al., 2014b). In the course of our investigation into how Tma20 and Tma22 facilitate uORF mediated repression of *BRE4* we showed that they also decrease the transcript levels of an allele of *BRE4* that we created which contains an oORF (Figure 30). Since Tma20 and Tma22 also facilitate the oORF mediated repression of *Stn1* we suggest that Tma20 and Tma22 may function to repress the translation of other genes that contain oORFs. In contrast with *BRE4* however our data suggests Tma20 and Tma22 decrease expression of *STN1* by decreasing translation, although it remains to be tested if Tma20 and Tma22 decrease translation re-initiation (Figure 36). It would be interesting to introduce a stop codon mutation of *STN1* oORF into to test whether translation re-initiation is mediated by the ribosome shifting in a 3' to 5' direction. We suggest a model whereby interaction of Tma20 and Tma22 with a uORF or oORF can have multiple possible consequences for expression of the corresponding main ORF (Figure 32).

One thing that is apparent from our analysis is that the interactions between Tma20 and Tma22 and uORF are rare and that most genes that have uORFs are unaffected by Tma20 and Tma22. It is unclear what factors cause Tma20 and Tma22 to regulate the expression of some genes that contain uORFs but not others.

*Drosophila* and mammalian homologs of Tma20 and Tma22 regulate expression of short uORFs in strong initiation contexts (Schleich et al., 2014b) (Schleich et al., 2017). 18 of the 22 uORF containing genes that we used to measure regulation by Tma20 and Tma22 (shown in Figure 25), have uORFs that are in stronger initiation contexts than *BRE4* therefore it is unlikely that the uORF initiation context is a factor which causes Tma20 and Tma22 dependence. This is in agreement with our observations that uORF initiation contexts do not play a major role in contributing to the repressive effects of uORFs (Figure 22). It is possible that uORF length or proximity to the main ORF are factors that determine if Tma20 and Tma22 affect gene expression. This is supported by the fact that *BRE4* and *STN1* both have a relatively short uORF and oORF respectively both of which are in close proximity to their corresponding main ORFs. However there are many of the other cases in Figure 25 that have uORFs shorter, and in closer proximity to the main ORF than the

*BRE4* uORF suggesting that having a short uORF or a uORF that is close to the main ORF does not impart Tma20 and Tma22 dependence. It would be necessary the regulation of more genes by Tma20 and Tma22 to determine what features of the uORF or indeed of the TL impart Tma20 and Tma22 dependence.



## Appendix A. Strains

DYL	Genotype	Figure
3001	<i>MATalpha ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 GAL+ psi+ ssd1-d2 RAD5</i>	11, 12, 13, 14, 15, 16, 17, 18, 19, 24, 28, 29, 30, 31, 33, 35, 36, 37, 38, 39
8460	<i>MATa ade2-1 trp1-1 can1-100 leu2-3,112 his3-11,15 ura3 GAL+ psi+ ssd1-d2 RAD5</i>	24, 28, 29, 30, 35, 37
1108	<i>MATa cdc13-1</i>	11, 12, 13, 14, 15, 16, 17, 18, 19
7026	<i>MATa cdc13-1 tma20::KANMX</i>	11, 13, 14, 17, 18, 19
8538	<i>MATalpha cdc13-1 tma22::KANMX</i>	11, 13, 15, 16, 17
8655	<i>MATa cdc13-1 tma64::NATMX</i>	11
8717	<i>MATa cdc13-1 tma22::HPHMX tma20::KANMX tma64::NATMX</i>	11
8566	<i>MATa cdc13-1 tma22::HPHMX</i>	11
8993	<i>MATa stn1-13</i>	11
8994	<i>MATa stn1-13</i>	11
9762	<i>MATalpha stn1-13 tma20::KANMX</i>	11
9763	<i>MATa stn1-13 tma20::KANMX</i>	11
9765	<i>MATa stn1-13 tma22::HPH</i>	11
9766	<i>MATa stn1-13 tma22::HPH</i>	11
9768	<i>MATa stn1-13 tma22::HPH tma20::KANMX</i>	11
1412	<i>MATa yku70::HIS3</i>	11
11933	<i>MATa yku70::HIS3</i>	11
11939	<i>MATalpha ku70::HIS3 tma20::KANMX</i>	11
11943	<i>MATalpha ku70::HIS3 tma20::KANMX</i>	11
4528	<i>MATa nmd2::HIS3</i>	12, 30, 36
6866	<i>MATa nmd2::HIS3</i>	12
8528	<i>MATa tma20::KANMX</i>	12, 24, 28, 30, 31, 36, 37

8529	<i>MATalpha tma20::KANMX</i>	12, 37
8542	<i>MATalpha tma22::HPH</i>	28, 31
8640	<i>MATa tma20::KANMX tma22::HPH</i>	28
8658	<i>MATa tma64::NATMX</i>	28
8623	<i>MATa tma20::KANMX nmd2::HIS3</i>	30, 36
11361	<i>MATa cdc13-1 HCR1::HIS3</i>	13,
11362	<i>MATa cdc13-1 HCR1::HIS3</i>	13, 14
11369	<i>MATa cdc13-1 tma22::HPH HCR1::HIS3</i>	13
11370	<i>MATalpha cdc13-1 tma20::KANMX</i>	13
11363	<i>MATa cdc13-1 tma20::kanmx HCR1::HIS3</i>	13
11364	<i>MATalpha cdc13-1 tma20::kanmx HCR1::HIS3</i>	13
11365	<i>MATalpha cdc13-1 tma22::HPH HCR1::HIS3</i>	13
11366	<i>MATa cdc13-1 tma22::HPH HCR1::HIS3</i>	13
11367	<i>MATalpha cdc13-1 tma22::HPH HCR1::HIS3</i>	13
11368	<i>MATa cdc13-1 tma22::HPH HCR1::HIS3</i>	14
11372	<i>MATa hcr1::HIS3</i>	14
5106	<i>MATa cdc13-1 nmd2::HIS3</i>	15
8619	<i>MATalpha cdc13-1-int nmd2::HIS3 tma20::KANMX</i>	15
8670	<i>MATa cdc13-1 nmd2::HIS3 tma22::KANMX</i>	15
4576	<i>MATa cdc13-1 ebs1::KANMX</i>	15
8627	<i>MATalpha cdc13-1-int ebs1::URA3 tma20::KANMX</i>	15
8633	<i>MATalpha cdc13-1-int ebs1::URA3 tma22::KANMX</i>	15
1289	<i>MATalpha cdc13-1 chk1::HIS3</i>	17
8729	<i>MATalpha cdc13-1-int tma20::KANMX chk1::HIS3</i>	17
8730	<i>MATalpha cdc13-1-int tma22::HPHMX chk1::HIS3</i>	17

1255	<i>MATa cdc13-1 rad9::HIS3</i>	17
8513	<i>MATa cdc13-1-int rad9::HIS3 tma20::KANMX</i>	17
8540	<i>MATa cdc13-1 tma22::KANMX rad9::HIS3</i>	17
1257	<i>MATa cdc13-1 rad24::TRP1</i>	17
8516	<i>MATa cdc13-1 rad24::TRP1 tma20::KANMX</i>	17
8687	<i>MATa cdc13-1-int rad24::TRP1 tma22::KANMX</i>	17
1296	<i>MATa cdc13-1 exo1::LEU2</i>	17
8517	<i>MATalpha cdc13-1 exo1::LEU2 tma20::KANMX</i>	17
8690	<i>MATa cdc13-1-int exo1::LEU2 tma22::KANMX</i>	17
1259	<i>MATa cdc13-1 rad9::HIS3 rad24::TRP1</i>	17
8521	<i>MATalpha cdc13-1-int rad9::HIS3 rad24::TRP1 tma20::KANMX</i>	17
8691	<i>MATalpha cdc13-1-int rad24::TRP1 rad9::HIS3 tma22::KANMX</i>	17
1692	<i>MATa cdc13-1 exo1::LEU2 rad9::HIS3</i>	17
8522	<i>MATa cdc13-1 rad9::HIS3 exo1::LEU2 tma20::KANMX</i>	17
8693	<i>MATalpha cdc13-1 exo1::LEU2 rad9::HIS3 tma22::KANMX</i>	17
1696	<i>MATalpha cdc13-1 exo1::LEU2 rad24::TRP1</i>	17
8525	<i>MATa cdc13-1 rad24::TRP1 exo1::LEU2 tma20::KANMX</i>	17
8696	<i>MATa cdc13-1 exo1::LEU2 rad24::TRP1 tma22::KANMX</i>	17
1694	<i>MATalpha cdc13-1 exo1::LEU2 rad9::HIS3 rad24::TRP1</i>	17
8526	<i>MATa cdc13-1 rad24::TRP1 rad9::HIS3 exo1::LEU2 tma20::KANMX</i>	17
8697	<i>MATalpha cdc13-1 exo1::LEU2 rad24::TRP1 rad9::HIS3 tma22::KANMX</i>	17

7023	<i>MATalpha cdc13-1 tma20::KANMX rad9::HIS3</i>	16
7028	<i>MATalpha cdc13-1 tma20::KANMX</i>	18
10714	<i>MATa cdc13-1 TMA20-3FLAG-KANMX</i>	18
10404	<i>MATa cdc13-1 TMA20-3FLAG-KANMX</i>	18
10407	<i>MATa TMA20-3FLAG-KANMX</i>	18
10406	<i>MATalpha TMA20-3FLAG-KANMX</i>	18
12287	<i>MATalpha STN1-URA3 cdc13-1</i>	35, 38
12301	<i>MATa URA3-STN1-102 cdc13-1</i>	35, 38
12297	<i>MATalpha URA3-STN1-102 cdc13-1</i>	35
12317	<i>MATa URA3-STN1-101 cdc13-1</i>	35, 38
12313	<i>MATalpha URA3-STN1-101 cdc13-</i>	35
11817	<i>MATa nmd2::HIS STN1-MYC</i>	36
11819	<i>MATalpha nmd2::HIS STN1-MYC</i>	36
11827	<i>MATalpha tma20::KANMX STN1-MYC</i>	36
11828	<i>MATalpha tma20::KANMX STN1-MYC</i>	36
11823	<i>MATa nmd2::HIS tma20::KANMX STN1-MYC</i>	36
11824	<i>MATa nmd2::HIS tma20::KANMX STN1-MYC</i>	36
12294	<i>MATa URA3-STN1</i>	39
12309	<i>MATa URA3-STN1-102</i>	39
12305	<i>MATalpha URA3-STN1-102</i>	39
12321	<i>MATalpha URA3-STN1-101</i>	39
12325	<i>MATa URA3-STN1-101</i>	39
10876	<i>MATa rad50::URA3</i>	39
11813	<i>MATa STN1-MYC</i>	34, 36
11814	<i>MATa STN1-MYC</i>	34, 36
12049	<i>MATalpha stn1-uAUG1-1-MYC-TRP</i>	34
12050	<i>MATalpha stn1-uAUG1-1-MYC-TRP</i>	34

12051	<i>MATalpha stn1-uAUG1-1-MYC-TRP</i>	34
12060	<i>MATalpha stn1-uAUG2-1-MYC-TRP</i>	34
12061	<i>MATalpha stn1-uAUG2-1-MYC-TRP</i>	34
12062	<i>MATalpha stn1-uAUG2-1-MYC-TRP</i>	34
12289	<i>MATalpha URA3-STN1 cdc13-1 tma20::KANMX</i>	38
12302	<i>MATa URA3-STN1-102 cdc13-1 tma20::KANMX</i>	38
12314	<i>MATalpha URA3-STN1-101 cdc13-1 tma20::KANMX</i>	38
12292	<i>MATa URA3-STN1 cdc13-1 tma20::KANMX</i>	38
12299	<i>MATalpha URA3-STN1-102 cdc13-1 nmd2::HIS3</i>	38
12319	<i>MATa URA3-STN1-101 cdc13-1 nmd2::HIS3</i>	38

## Appendix B. Plasmids

### Plasmids obtained from Lydall lab collection and others

<i>pDL</i>	<i>Description</i>	<i>Source/Reference</i>
1674	<i>SUP35NM fused to GFP under a CUP1 URS</i>	<i>Mick Tuite (p6442)</i>
1277	<i>URA3 centromeric plasmid (pRS416)</i>	<i>pDL13; pRS416 genetics 1989 122:19-27</i>
452	<i>pRS406</i>	<i>genetics 1989 122:19-27</i>
1833	<i>pFA6 URA3</i>	<i>Houseley et al (NAR) 2011</i>
1250	<i>LEU centromeric plasmid (pRS415)</i>	<i>pDL13; pRS415 genetics 1989 122:19-27</i>
1659	<i>plasmid containing Renilla and firefly Luciferase</i>	<i>Keeling et al 2004</i>
452	<i>pRS306</i>	<i>pRS306 genetics 1989 122:19-27</i>
216	<i>pRS425</i>	<i>genetics 1989 122:19-27</i>

### Plasmids created in this study by *in vivo* cloning

<i>pDL</i>	<i>Description</i>	<i>Fragments</i>	<i>Primers</i>
1798	<i>TMA20 expressed from CEN plasmid (LEU)</i>	<i>TMA20 ORF</i>	<i>m3542 and m3566 from DLY 3001 DNA</i>
	<i>backbone - pDL1250 digested with BAMH1 BRE4</i>		
1799	<i>TMA20 S118A T119A expressed from CEN plasmid (LEU).</i>	<i>Frag 1 - TMA20 S118A T119A Reporter</i>	<i>Part 1 amplified from pDL1798 with M1918 and M4251</i>
		<i>Frag 2 - TMA20 S118A T119A Reporter</i>	<i>Part 2 amplified from pDL1798 with M1919 and M4252</i>
		<i>backbone - pDL1250 digested with BAMH1</i>	

1665	<i>TMA20 expressed from CEN plasmid (URA3)</i>	<i>TMA20 ORF</i>	<i>m3542 and m3566 from DLY 3001 DNA</i>
		<i>backbone - pDL1277 digested with BAMH1</i>	
1676	<i>TMA20 T79A - expressed from CEN plasmid (URA3)</i>	<i>Frag 1 - T79A Reporter</i>	<i>m1918 and m3599 amplified from pDL1665</i>
		<i>Frag 2 - T79A Reporter</i>	<i>m1919 and m3599 amplified from pDL1665</i>
		<i>backbone - pDL1277 digested with BAMH1</i>	
1701	<i>pRS415 with PGK1 promoting Firefly</i>	<i>Frag1 - PGK1 promoter</i>	<i>m3926 and m3812 from DLY 3001</i>
		<i>Frag 2 - Firefly ORF</i>	<i>m3829 and m3830</i>
1728	<i>PGK1 drives Firefly luciferase YFG drives Renilla luciferase</i>	<i>CYC1 terminator sequence</i>	<i>m3962 and m3965 from DLY 3001 DNA</i>
		<i>ADH1 terminator</i>	<i>m3962 and m3963 from DLY 3001 DNA</i>
		<i>Renilla luciferase</i>	<i>m3966 and m3964 from pDL1659</i>
		<i>backbone - pDL1701 digested with Spe1 and BamH1</i>	
1777	<i>CTF3 Reporter</i>	<i>CTF3 URS</i>	<i>m4147 and m4148 from DLY 3001 DNA</i>
1788	<i>CTF3-M1V Reporter</i>	<i>Frag 1 - CTF3 M1V Reporter</i>	<i>m4035 and m4188 used to amplify part1 from pDL1777</i>
		<i>Frag 2 - CTF3 M1V Reporter</i>	<i>m4038 and m4187 used to amplify part 2 from pDL1777</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1767	<i>BRE4 Reporter</i>	<i>BRE4 URS</i>	<i>m4169 and m4170 from DLY 3001 DNA</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1768	<i>BRE4-9*Y Reporter</i>	<i>Frag 1- BRE4-9*Y URS</i>	<i>m4038 and m4230 from pDL 1767</i>

		<i>Frag 1- BRE4-9*Y URS</i>	<i>m4035 and m4231 from pDL 1767</i>
1766	<i>HDA3 Reporter</i>	<i>HDA3 URS</i>	<i>m4167 and m4168 amplified from DLY 3001</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1768	<i>HDA3 URS with non functional uORFs inserted in front of Renilla in pDL1728</i>	<i>Frag1 - HDA3 M1V Reporter</i>	<i>m4035 and m4203 used to amplify part1 from pDL1766</i>
		<i>Frag2 - HDA3 M1V Reporter</i>	<i>m4038 and m4205 used to amplify part 2 from pDL1766</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1782	<i>NTR2 Reporter</i>	<i>NTR2 URS</i>	<i>m4157 and m4162 from DLY 3001 DNA</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1793	<i>NTR2-M1V Reporter</i>	<i>Frag 1 - NTR2 M1V Reporter</i>	<i>m4035 and m4202 used to amplify part1 from pDL1782</i>
		<i>Frag 2 - NTR2 M1V Reporter</i>	<i>m4038 and m4203 used to amplify part 2 from pDL1782</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1781	<i>ATG5 Reporter</i>	<i>ATG5 URS</i>	<i>m4155 and m4156 from DLY 3001</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1792	<i>ATG5-M1V Reporter</i>	<i>Frag 1 - ATG5 M1V Reporter</i>	<i>m4035 and m4182 used to amplify part 1 from pDL1781</i>
		<i>Frag 2 - ATG5 M1V Reporter</i>	<i>m4038 and m4183 used to amplify part 2 from pDL1781</i>
		<i>backbone - pDL1728 digested with</i>	



		<i>Not1</i>	
1780	<i>NDJ1 Reporter</i>	<i>NDJ1 URS</i>	<i>m4153 and m4154 from DLY 3001 DNA</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1791	<i>NDJ1 M1V Reporter</i>	<i>Frag 1 - NDJ1 M1V URS</i>	<i>m4035 and m4193 used to amplify part1 from pDL1780</i>
		<i>Frag 2 - NDJ1 M1V URS</i>	<i>m4038 and m4194 used to amplify part 2 from pDL1780</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1745	<i>PGS1 Reporter</i>	<i>PGS1 URS</i>	<i>m4009 and m4010 from DLY 3001 DNA</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1840	<i>PGS1-M1V Reporter</i>	<i>Frag 1 - PGS1 M1V URS</i>	<i>part1 amplified using m4035 and m4132 with pDL1745</i>
		<i>Frag 2 - PGS1 M1V URS</i>	<i>part2 amplified using m4038 and m4133 with pDL1745</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1778	<i>FRE6 Reporter</i>	<i>FRE6 URS inserted in front of Renilla in pDL1728</i>	<i>m4149 and m4150 from DLY 3001</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1790	<i>FRE6-M1V Reporter</i>	<i>Frag 1 - FRE6 M1V URS</i>	<i>m4038 and m4189 used to amplify part1 from pDL1778</i>
		<i>Frag 2 - FRE6 M1V URS</i>	<i>m4035 and m4190 used to amplify part 2 from pDL1778</i>
		<i>backbone - pDL1728 digested with Not1</i>	

1783	RRG1 Reporter	RRG1 URS	m4158 and m4159 from DLY 3001 DNA
		backbone - pDL1728 digested with Not1	
1794	RRG1-M1V Reporter	Frag 1 - RRG1 M1V URS	m4035 and m4213 used to amplify part1 from pDL1783
		Frag 2 - RRG1 M1V URS	m4038 and m4214 used to amplify part 2 from pDL1783
		backbone - pDL1728 digested with Not1	
1779	PUS9 Reporter	PUS9 URS inserted in front of Renilla in pDL1728	m4151 and m4152 from DLY 3001 DNA
		backbone - pDL1728 digested with Not1	
1789	PUS9-M1V Reporter	Frag 1 - PUS9 M1V URS	m4035 and m4181 used to amplify part1 from pDL1779
		Frag 2 - PUS9 M1V URS	m4038 and m4182 used to amplify part 2 from pDL1779
		backbone - pDL1728 digested with Not1	
1743	MOD5 Reporter	MOD5 URS	m4019 and m4020 from DLY 3001
		backbone - pDL1728 digested with Not1	
1843	MOD5-M1V Reporter	Frag 1 - MOD5 M1V URS	m4035 and m4123 used to amplify part1 from pDL1743
		Frag 2 - MOD5 M1V URS	m4024 and m4125 used to amplify part2 from pDL1743
		Frag 2 - MOD5 M1V URS	m4038 and m4130 used to amplify part 3 from pDL1743

1776	ATG20 Reporter	ATG20 URS	m4145 and m4146 from DLY 3001
		backbone - pDL1728 digested with Not1	
1787	ATG20-M1V Reporter	Frag 1 - ATG20 M1V URS	m4035 and m4186 used to amplify part1 from pDL1776
		Frag 2 - ATG20 M1V URS	m4038 and m4185 used to amplify part 2 from pDL1776
1749	DUS3 Reporter	DUS3 URS	m4021 and m4022 from DLY 3001
		backbone - pDL1728 digested with Not1	
1841	DUS3-M1V Reporter	Frag 1 - DUS3 M1V URS	m4035 and m4138 used to amplify part1 from pDL1749
		Frag 2 - DUS3 M1V URS	m4038 and m4139 used to amplify part 2 from pDL1749
		backbone - pDL1728 digested with Not1	
1785	SMY1 Reporter	SMY1 URS	m4171 and m4172 from DLY 3001
		backbone - pDL1728 digested with Not1	
1796	SMY1-M1V Reporter	Frag 1 - SMY1 M1V URS	m4035 and m4198 used to amplify part1 from pDL1785
		Frag 2 - SMY1 M1V URS	m4038 and m4197 used to amplify part 2 from pDL1785
1786	MBP1 Reporter	MBP1 URS inserted in front of Renilla in pDL1728	m4173 and m4174 from DLY 3001
		backbone - pDL1728 digested with Not1	

1797	MBP1-M1V Reporter	Frag 1 - MBP1 M1V URS	m4035 and m4200 used to amplify part1 from pDL1786
		Frag 2 - MBP1 M1V URS	m4038 and m4199 used to amplify part 2 from pDL1786
		backbone - pDL1728 digested with Not1	
1746	RBS1 Reporter	RBS1 URS inserted in front of Renilla in pDL1728	m4011 and m4012
		backbone - pDL1728 digested with Not1	
1754	RBS1-M1V Reporter	Frag 1 - RBS1 M1V URS	m4035 and m4041 used to amplify part1 from pDL1746
		Frag 2 - RBS1 M1V URS	m4038 and m4042 used to amplify part 2 from pDL1746
		backbone - pDL1728 digested with Not1	
1747	SFG1 Reporter	SFG1 URS inserted in front of Renilla in pDL1728	m4015 and m4016 from DLY 3001
		backbone - pDL1728 digested with Not1	
1753	SFG1-M1V Reporter	Frag 1 -SFG1 M1V URS	m4035and m4039 used to amplify part1 from pDL1747
		Frag 2 - SFG1 M1V URS	m4038 and m4040 used to amplify part 2 from pDL1747
		backbone - pDL1728 digested with Not1	
1750	CDC13 Reporter	CDC13 URS inserted in front of Renilla in pDL1728	m4023 and m4024 from DLY 3001
		backbone - pDL1728 digested with Not1	
1888	CDC13 M1V	Frag 1 - CDC13 M1V URS	m4035 and m4127 from

	<i>Reporter</i>		<i>pDL1750</i>
		<i>Frag 2 - CDC13 M1V URS</i>	<i>m4131 and m4038 from pDL1750</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1784	<i>IST2 Reporter</i>	<i>IST2 URS inserted in front of Renilla in pDL1728</i>	<i>m4160 and m4161</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1795	<i>IST2-M1V Reporter</i>	<i>Frag 1 -IST2 M1V URS</i>	<i>m4035 and m4192 used to amplify part1 from pDL1784</i>
		<i>Frag 2 - IST2 M1V URS</i>	<i>m4038 and m4191 used to amplify part 2 from pDL1784</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1748	<i>RAI1 Reporter</i>	<i>RAI1 URS inserted in front of Renilla in pDL1728</i>	<i>m4017 and m4018</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1842	<i>RAI1-M1V Reporter</i>	<i>Frag 1 -RAI1 M1V URS</i>	<i>m4035 and m4134 used to amplify part1 from pDL1748</i>
		<i>Frag 2 - RAI1 M1V URS</i>	<i>m4038 and m4135 used to amplify part 2 from pDL1748</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1744	<i>CMK1 Reporter</i>	<i>CMK1 URS inserted in front of Renilla in pDL1728</i>	<i>m4007 and m4008</i>
		<i>backbone - pDL1728 digested with Not1</i>	
1844	<i>CMK1-M1V Reporter</i>	<i>Frag 1 - CMK1 M1V URS</i>	<i>m4035 and m4136 used to amplify part1 from pDL1744</i>
		<i>Frag 2 - CMK1 M1V URS</i>	<i>m4038 and m4137 used</i>

			to amplify part 2 from pDL1744
		backbone - pDL1728 digested with Not1	
1743	ALD3 Reporter	ALD3 URS	m4019 m4020
		backbone - pDL1728 digested with Not1	
1843	ALD3- M1V URS	Frag 1 - ALD3 M1V URS	m4035 and m4123 from pDL1745
		Frag 2 - ALD3 M1V URS	m4124 and m4125 from pDL1745
		Frag 3 - ALD3 M1V URS	m4038 and m4127 from pDL1745
		backbone - pDL1728 digested with Not1	
1895	RLI1 expressed in 2micron plasmid (leu2)	RLI1 ORF	m4384 and m4385 with DLY 3001
		prs425 (pDL216) digested with BamH1	
1889	ALD3 uORF1 M1V	Frag 1 - ALD3 uORF 1M1V URS	m4035 and m4277 from pDL1745
		Frag 2 - ALD3 uORF 1M1V URS	m4038 and m4278 <b>from pDL1745</b>
		backbone - pDL1728 digested with Not1	
1890	ALD3 uORF2 M1V	Frag 1 - ALD3 uORF2 1M1V URS	m4035 and m4279 from pDL1745
		Frag 2 - ALD3 uORF2 1M1V URS	m4038 and m4280 from pDL1745
		backbone - pDL1728 digested with Not1	
1891	ALD3 uORF3 M1V	Frag 1 - ALD3 uORF3 1M1V URS	m4035 and m4281 from pDL1745
		Frag 2 - ALD3 uORF3 1M1V URS	m4038 and m4282 from pDL1745
		backbone - pDL1728 digested with Not1	

1893	BRE4 uORF 9*Y	Frag 1 - BRE4 uORF 9*Y URS	m4035 and m4231 from pDL1767
		Frag 2 - BRE4 uORF 9*Y URS	m4038 and m4230 from pDL1767
		backbone - pDL1728 digested with Not1	
1894	BRE4 out	Frag 1 - BRE4 out URS	m4035 and m4380 from pDL1767
		Frag 2 - BRE4 out URS	m4038 and m4379 from pDL1767
		backbone - pDL1728 digested with Not1	

### Plasmids created during this study by Gibson assembly

1867	Integrate STN1 with URA3 in the URS into genome	Frag 1 - STN1 URS and first 139 bps of STN1	m4608 + m4609 from 3001 DNA
		Frag 2 - URA3 in including endogenous URS and terminator in the same orientation as STN1	m4610 + m4611 from pDL1833
		Frag 3 - PDC2 URS	m4612 + m4613 from 3001 DNA
		pDL 452 digested with BamH1 and Xho1	
1868	Integrate STN1-102 with URA3 in the URS into genome	Frag 1 - STN1 URS and first 139 bps of STN1	m4608 + m4609 from 11870 DNA
		Frag 2 - URA3 in including endogenous URS and terminator in the same orientation as STN1	m4610 + m4611 from pDL1833
		Frag 3 - PDC2 URS	m4612 + m4613 from 3001 DNA
		backbone - pDL452 (cut with BamH1 and Xho1)	
1869	Integrate STN1-101 with URA3 in the URS into genome	Frag 1 - STN1 URS and first 139 bps of STN1	m4608 + m4609 from 11871 DNA
		Frag 2 - URA3, including endogenous URS and terminator	m4610 + m4611 from pDL1833
		Frag 3 - PDC2 URS	m4612 + m4613 from 3001 DNA

		backbone - pDL452 (cut with BamH1 and Xho1)	
--	--	---	--



## Appendix C. Primer sequences

Primer	Sequence
m3599	CGATGAACTGATTCCAGTCCTGAAATTAGTACACAAATTTCC
m3600	CTGGGACTTACCTGTCCTGA
m3542	AACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCAT ACGCTTCACGAAGAACG
m3566	CAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGGCT AGGTTTGATCGTGTGATGG
m4251	GCGGTGTGAAATACCGCACAGATGCGTAAGG
m4252	CTGGAAAGCGGGCAGTGAGCGCAACGC
m4147	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGCCA AGCCACGCGACG
m4148	TCCGTTTCCTTTGTTCTGGATCATAAACTTTTGAAGTCATTCATG TAAAAAAAACAAAAATTGA
m4035	GGTCTGGTATAATACACCGCGCTAC
m4188	ATCAAAAACAATCAACTGTGAAG
m4038	CCTATGTTGTGTGGAATTGTGAGCGG
m4187	CGTATTGTTTTTCACTTCACAGTTG
m4169	TCCGTTTCCTTTGTTCTGGATCATAAACTTTTGAAGTCATGAGG ACCTGTTACAATTCCAG
m4170	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGGCA AACTACTACTAACAAC
m4167	TCCGTTTCCTTTGTTCTGGATCATAAACTTTTGAAGTCATACTAA AGGGCAATTAGTTTGTG
m4168	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGAGTT ACTGATTTAATCCACTCAG
m4203	GAAATACTTTAGACTATAAAAGTG
m4157	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTATTA ATTACTGCAAACCCAA
m4162	TCCGTTTCCTTTGTTCTGGATCATAAACTTTTGAAGTCATAATGT GTTCAAATGCTATAAC
m4202	GTTACGTCAAGTTCTGCCTAGCAGTG

m4203	GAAATACTTTAGACTATAAAAAGTG
m4155	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTAC CGGATAACAGAGTAA
m4156	TCCGTTTCCTTTGTTCTGGATCATAACTTTCTGAAGTCATACTTA CATCATAGGTTTCCT
m4182	CATAAGTTTCCTATTCCACTATACGAGCACTCACCC
m4183	CTTCTAGAACCACAAAAGAACAAAATTCACTCGCACTTCTGAGCA CGCTACTTCACTATTT
m4153	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGAATGT ATTACCTGACTCAGG
m4154	TCCGTTTCCTTTGTTCTGGATCATAACTTTCTGAAGTCATTATAG TTTTTCTACCTGTATTTT
m4194	GTATGCGAATTAAAGTAATCCGTGTG
m4193	TGATTTTATTTACCACACGGA
m4205	CTTCTAACTTTCTTTTTATCACTTTTATAG
m4009	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTTATT GCAATTACTTCTTCTCA
m4010	TCCGTTTCCTTTGTTCTGGATCATAACTTTCTGAAGTCATTAATA TATGTTATCCTGAGTAT
m4132	CACAAGAAAGTAGATATAGTGTAGG
m4133	GGACAAGCTGGGTGTCCTACACTATATC
m4149	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGGTA AATGATTTACAAATGA
m4150	TCCGTTTCCTTTGTTCTGGATCATAACTTTCTGAAGTCATATTCA ACGAGCTATGGTC
m4189	CGTCACGCTTCACTTTAGAAGATTATAGAAAAATCACTGC
m4190	GCAGTGATTTTTCTATAATCTTCTAAAGTGAAGCGTGACG
m4158	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGAGTGT TTAAATCATTTATTTTCA
m4159	TCCGTTTCCTTTGTTCTGGATCATAACTTTCTGAAGTCATATTTT CTACTTGTTTATTACAA
m4213	AGATTAACACTTTTGTAGTG
m4214	GAAGTCATATTTTCTACTTGTTCACTACAAAAG
m4151	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGAGTTT

	TTGTATCGCGATGTT
m4152	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCGAAGTCATAAGTT TCCTATTCCATTATAC
m4181	GAAAATGGGTGAGTGCTCGTATAGTGG
m4182	CATAAGTTTCCTATTCCACTATACGAGCACTCACCC
m4019	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTTAGT AAAACCTAAAATATATGTAT
m4020	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCGAAGTCATGAAG ATTTTTGGTTCTATTTAC
m4123	ATTAAACAGAAAAAAATTCTAAAGTGGC
m4124	GCTTTTCTTCTTTTTTGCCACTTTAG
m4125	CCTTATGTTGCTGCTCCTCGGTGGC
m4127	CATATATGTTTCTCTTTGGATACGAGTGACC
m4145	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTTTGC TGGTAAAAAAAAGGTA
m4146	CCGTTTCCTTTGTTCTGGATCATAAACTTTCGAAGTCATTGTGT GCTGTAGAAGACAA
m4186	GATTTACAGCAAGTCATCTCCAAAGTGAAGAGGTG
m4185	CGAGCTAGTCACCTCTTCACTTTGG
m4021	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTAAGA TCTAGTTAAAGTTATGAA
m4022	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCGAAGTCATCAATA ATGTTTAGAAGTTGCC
m4138	GAGCAGTAATGAAACATTAGTGAAGTC
m4139	GCCTTTTTTTCCTATTCTGAGTTCACTAATG
m4171	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCGAAGTCATTTTAA CAAGTTTGCACCTAC
m4172	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTCGAT TTATAATACTAACTTAG
m4198	GGTTAACAAGGAAAGAGGTGGAAGTTAAAAAAGTGGAAT
m4197	ATTTCCACTTTTTTAACTTCCACCTCTTTCCTTGTTAACC
m4173	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCGAAGTCATGCTT GTGTTTCTGGGATTT
m4174	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGACCAT

	ACAACTACATATTAATC
m4200	CCATCCGATTAGTATAATCAACAGGTGAAAAAGTG
m4199	CTTTTTTTCACCTTTTTCACCTGTTGATTA
m4011	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTGATC TATGGAATGAGGGTT
m4012	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCTGAAGTCATCACC TACTTCTGCCCTAA
m4042	GTGCGCCCACTATAAGATGCACGTTTCAACACACTGTC
m4041	CCGTGGTGACAGTGTGTTGAAACGTGCATCTTATAGTGGGCG
m4015	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTAATG GATCAATTTTGTCACTC
m4016	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCTGAAGTCATGTTTT TGAAATTTTTTTTTCTTA
m4039	TGATAAGTTGTTCTGTAAATGTGTTTGATTAG
m4040	CTCTTGTTTAACTAATCAAACACATTTACA
m4023	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTTGTA AATCCGCTGCTGAAT
m4024	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCTGAAGTCATTTTAA GGCGATAGTTTCCAC
m4160	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGGTTTG TGCTGTACTAATCTA
m4161	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCTGAAGTCATAGTAT GCTCAACCGCTTG
m4192	GGTGAATACTGTTTCGGTCGATTTGTGG
m4191	GTATTTGACAAATTAGCCACAAATC
m4017	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTCTTG ATCCCTTCTATGCTT
m4018	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCTGAAGTCATTGGG ATACACTATGTTGGT
m4134	GTCTGGTAGTGTTACTAATACTTCTGTAATGTGG
m4135	CTTTCACCACATTACAGAAGTATTAGTAACACTA
m4007	CCCTCACTAAAGGGAACAAAAGCTGGAGCTCCACCGCGTAACC ATCTTAACTTCTGCTA
m4008	TCCGTTTCCTTTGTTCTGGATCATAAACTTTCTGAAGTCATAGTAA

	TTGACTCGTTATTTTTT
m4136	CTACAAGCCAGAGTATAAGTGG
m4137	GAGTTCTCTACGAGAATTGAAACCACTTATAC
m4384	AACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCGAAT ATTTGACTACTTTGTTTGG
m4385	CAAGCGCGCAATTAACCCTCACTAAAGGGAACAAAAGCTGCTG TTGCTATAGTAGCTAAAGCGG
m4277	CATACCTCAGAACTAATTGTATAAATTGGC
m4278	GTCAAATTTATTCCCATCCAGATTTGCCAATTTATAC
m4279	CTAATTGTATAAAATGGCAAATCTGGTTGGG
m4280	GTCAAATTTATTCCCAACCAGATTTGCC
m4281	GGGAATAAATTTGACTCTTGCCAGG
m4282	CCAAGGTATAATTACCTGGCAAGAG
m4231	GTATAGTGCTGGAATTGTATCAGG
m4230	CTTTCGAAGTCATGAGGACCTGATAC
m4380	CAGTATAGTGCTGGAATTGGTAAC
m4379	CGAAGTCATGAGGACCTGTTACCAATTCC
m4609	CGAAGTCATGAGGACCTGTTACCAATTCC
m4611	TAACACCAAGCAGTAAAGAGACAGCTTTATTATAACCAGCGAAT TCGAGCTCGTTTAAAC
m4613	GCTTTATGGAGGATCTGGCGCGCCTTAATTAACCCGGGGATCC GGCTCCAACCCATTATTAAAAG
m4610	AAAGAGCAATAAGCTGGCTTTTAATAATGGGTTGGAGCCGGAT CCCCGGGTAAATTAAG
m4608	ATATCAGTTATTACCCGGGCTGTTTAAACGAGCTCGAATTCGCT GGTTATAATAAGCTGTCTC
m4612	GGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGTG CCAAACAGCTTGTGCAG
m4609	CAAGCTCGGAATTAACCCTCACTAAAGGGAACAAAAGCTGCTC TCATCCGGTACTTTAAATC

### Primers for gene deletion

	Sequence	Target
m2858	TTCAAGAGCTAAACTAAAGAAAAGCATATTGCAT AAAATGCGGATCCCCGGGTAAATTAA	TMA22
m2862	TCTCCCAGAACGGTGCTATTACATATTTATGGAT TGCTTAGAATTCGAGCTCGTTTAAAC	TMA22
m4296	CTATCCTAACCACCACCTCAAAAAAAAAAAGTA ATAAAACGGATCCCCGGGTAAATTAA	HCR1
m4298	AGATGGACAAGTTTATCATAGCAAAGAAACAATA AGCAGAGAATTCGAGCTCGTTTAAAC	HCR1
m3894	AGATGGACAAGTTTATCATAGCAAAGAAACAATA AGCAGAGAATTCGAGCTCGTTTAAAC	TMA20-flag
m3896	GCAGATGGATAGTAATATAGTGTTGACGGCTCC GTTTG GAATTCGAGCTCGTTTAAAC	TMA20-flag
m2922	CCGACTCAATAGATTAGTGTAGCGCAGGATTAG TACAGCTCGGATCCCCGGGTAAATTAA	TMA64
m2923	CGGGCATTTTTACGCATTTAAACATTTATATGATA TAAATGAATTCGAGCTCGTTTAAAC	TMA64

### Primers for qPCR

	Sequence	Target
m1172	CAGACCGAACTCGGTGATT T	BUD6
m1173	TTTTAGCGGGCTGAGACCT A	BUD6
m1734	TCGAGCAACTGCAAGAAGA A	STN1
m1735	CGAAATGACAAGGAATGCA C	STN1

## References

- ADAMS MARTIN, A., DIONNE, I., WELLINGER, R. J. & HOLM, C. 2000. The function of DNA polymerase alpha at telomeric G tails is important for telomere homeostasis. *Mol Cell Biol*, 20, 786-96.
- ADDINALL, S. G., DOWNEY, M., YU, M., ZUBKO, M. K., DEWAR, J., LEAKE, A., HALLINAN, J., SHAW, O., JAMES, K. & WILKINSON, D. J. 2008. A genomewide suppressor and enhancer analysis of *cdc13-1* reveals varied cellular processes influencing telomere capping in *Saccharomyces cerevisiae*. *Genetics*, 180, 2251-2266.
- ADDINALL, S. G., HOLSTEIN, E. M., LAWLESS, C., YU, M., CHAPMAN, K., BANKS, A. P., NGO, H. P., MARINGELE, L., TASCHUK, M., YOUNG, A., CIESIOLKA, A., LISTER, A. L., WIPAT, A., WILKINSON, D. J. & LYDALL, D. 2011. Quantitative fitness analysis shows that NMD proteins and many other protein complexes suppress or enhance distinct telomere cap defects. *PLoS Genet*, 7, e1001362.
- ADVANI, V. M., BELEW, A. T. & DINMAN, J. D. 2013. Yeast telomere maintenance is globally controlled by programmed ribosomal frameshifting and the nonsense-mediated mRNA decay pathway. *Translation (Austin)*, 1, e24418.
- ALGIRE, M. A., MAAG, D., SAVIO, P., ACKER, M. G., TARUN, S. Z., SACHS, A. B., ASANO, K., NIELSEN, K. H., OLSEN, D. S. & PHAN, L. 2002. Development and characterization of a reconstituted yeast translation initiation system. *Rna*, 8, 382-397.
- ALTAMURA, E., BORGATTI, M., FINOTTI, A., GASPARELLO, J., GAMBARI, R., SPINELLI, M., CASTALDO, R. & ALTAMURA, N. 2016. Chemical-Induced Read-Through at Premature Termination Codons Determined by a Rapid Dual-Fluorescence System Based on *S. cerevisiae*. *PLoS One*, 11, e0154260.
- ANBALAGAN, S., BONETTI, D., LUCCHINI, G. & LONGHESE, M. P. 2011. Rif1 supports the function of the CST complex in yeast telomere capping. *PLoS Genet*, 7, e1002024.
- ANDREEV, D. E., O'CONNOR, P. B., FAHEY, C., KENNY, E. M., TERENCE, I. M., DMITRIEV, S. E., CORMICAN, P., MORRIS, D. W., SHATSKY, I. N. & BARANOV, P. V. 2015. Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. *Elife*, 4, e03971.
- ARAVA, Y., WANG, Y., STOREY, J. D., LIU, C. L., BROWN, P. O. & HERSCHLAG, D. 2003. Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*, 100, 3889-94.
- ARRIBERE, J. A. & GILBERT, W. V. 2013. Roles for transcript leaders in translation and mRNA decay revealed by transcript leader sequencing. *Genome Res*, 23, 977-87.
- BARBOSA, C., PEIXEIRO, I. & ROMAO, L. 2013a. Gene expression regulation by upstream open reading frames and human disease. *PLoS Genet*, 9, e1003529.
- BARBOSA, C., PEIXEIRO, I. & ROMÃO, L. 2013b. Gene expression regulation by upstream open reading frames and human disease. *PLoS genetics*, 9, e1003529.
- BECKER, T., FRANCKENBERG, S., WICKLES, S., SHOEMAKER, C. J., ANGER, A. M., ARMACHE, J. P., SIEBER, H., UNGEWICKELL, C., BERNINGHAUSEN, O., DABERKOW, I., KARCHER, A., THOMM, M., HOPFNER, K. P., GREEN, R. & BECKMANN, R. 2012. Structural basis of highly conserved ribosome recycling in eukaryotes and archaea. *Nature*, 482, 501-6.
- BELL, R. J., RUBE, H. T., XAVIER-MAGALHAES, A., COSTA, B. M., MANCINI, A., SONG, J. S. & COSTELLO, J. F. 2016. Understanding TERT Promoter Mutations: A Common Path to Immortality. *Mol Cancer Res*, 14, 315-23.
- BERTHELOT, K., MULDOON, M., RAJKOWITSCH, L., HUGHES, J. & MCCARTHY, J. E. G. 2004. Dynamics and processivity of 40S ribosome scanning on mRNA in yeast. *Molecular microbiology*, 51, 987-1001.
- BEZNOSKOVA, P., CUCHALOVA, L., WAGNER, S., SHOEMAKER, C. J., GUNISOVA, S., VON DER HAAR, T. & VALASEK, L. S. 2013. Translation initiation factors eIF3 and HCR1 control translation termination and stop codon read-through in yeast cells. *PLoS Genet*, 9, e1003962.
- BISIO, A., NASTI, S., JORDAN, J. J., GARGIULO, S., PASTORINO, L., PROVENZANI, A., QUATTRONE, A., QUEIROLO, P., BIANCHI-SCARRÀ, G. & GHIORZO, P. 2010. Functional analysis of

- CDKN2A/p16INK4a 5'-UTR variants predisposing to melanoma. *Human molecular genetics*, 19, 1479-1491.
- BORDEIRA-CARRICO, R., PEGO, A. P., SANTOS, M. & OLIVEIRA, C. 2012. Cancer syndromes and therapy by stop-codon readthrough. *Trends Mol Med*, 18, 667-78.
- BOULTON, S. J. & JACKSON, S. P. 1998. Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing. *EMBO J*, 17, 1819-28.
- BRAR, G. A., YASSOUR, M., FRIEDMAN, N., REGEV, A., INGOLIA, N. T. & WEISSMAN, J. S. 2012. High-resolution view of the yeast meiotic program revealed by ribosome profiling. *science*, 335, 552-557.
- CALVO, S. E., PAGLIARINI, D. J. & MOOTHA, V. K. 2009. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proceedings of the National Academy of Sciences*, 106, 7507-7512.
- CELIK, A., BAKER, R., HE, F. & JACOBSON, A. 2017a. High-resolution profiling of NMD targets in yeast reveals translational fidelity as a basis for substrate selection. *RNA*, 23, 735-748.
- CELIK, A., HE, F. & JACOBSON, A. 2017b. NMD monitors translational fidelity 24/7. *Curr Genet*.
- CERRUDO, C. S., GHIRINGHELLI, P. D. & GOMEZ, D. E. 2014. Protein universe containing a PUA RNA-binding domain. *FEBS Journal*, 281, 74-87.
- CHANG, D. T., HUANG, C. Y., WU, C. Y. & WU, W. S. 2011. YPA: an integrated repository of promoter features in *Saccharomyces cerevisiae*. *Nucleic Acids Res*, 39, D647-52.
- CHASTAIN, M., ZHOU, Q., SHIVA, O., WHITMORE, L., JIA, P., DAI, X., HUANG, C., FADRI-MOSKWIK, M., YE, P. & CHAI, W. 2016. Human CST Facilitates Genome-wide RAD51 Recruitment to GC-Rich Repetitive Sequences in Response to Replication Stress. *Cell Rep*, 16, 1300-14.
- CHEN, L.-Y., MAJERSKÁ, J. & LINGNER, J. 2013. Molecular basis of telomere syndrome caused by CTC1 mutations. *Genes & development*, 27, 2099-2108.
- CHEN, L.-Y., REDON, S. & LINGNER, J. 2012. The human CST complex is a terminator of telomerase activity. *Nature*.
- CHEN, Y., CALDWELL, J. M., PEREIRA, E., BAKER, R. W. & SANCHEZ, Y. 2009. ATRMec1 phosphorylation-independent activation of Chk1 in vivo. *Journal of Biological Chemistry*, 284, 182-190.
- CHEW, G. L., PAULI, A. & SCHIER, A. F. 2016. Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish. *Nat Commun*, 7, 11663.
- COSENTINO, G. P., SCHMELZLE, T., HAGHIGHAT, A., HELLIWELL, S. B., HALL, M. N. & SONENBERG, N. 2000. Eap1p, a novel eukaryotic translation initiation factor 4E-associated protein in *Saccharomyces cerevisiae*. *Molecular and cellular biology*, 20, 4604-4613.
- CROSS, F. R. 2015. Tying Down Loose Ends in the *Chlamydomonas* Genome: Functional Significance of Abundant Upstream Open Reading Frames. *G3 (Bethesda)*, 6, 435-46.
- DAHLSEID, J. N., LEW-SMITH, J., LELIVELT, M. J., ENOMOTO, S., FORD, A., DESRUISSEAU, M., MCCLELLAN, M., LUE, N., CULBERTSON, M. R. & BERMAN, J. 2003. mRNAs encoding telomerase components and regulators are controlled by UPF genes in *Saccharomyces cerevisiae*. *Eukaryot Cell*, 2, 134-42.
- DEVER, T. E. & GREEN, R. 2012. The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harb Perspect Biol*, 4, a013706.
- EAGLESTONE, S. S., RUDDOCK, L. W., COX, B. S. & TUIE, M. F. 2000. Guanidine hydrochloride blocks a critical step in the propagation of the prion-like determinant [PSI(+)] of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A*, 97, 240-4.
- ELLISON, V. & STILLMAN, B. 2003. Biochemical characterization of DNA damage checkpoint complexes: clamp loader and clamp complexes with specificity for 5' recessed DNA. *PLoS biology*, 1, e33.
- ENOMOTO, S., GLOWCZEWSKI, L., LEW-SMITH, J. & BERMAN, J. G. 2004. Telomere cap components influence the rate of senescence in telomerase-deficient yeast cells. *Mol Cell Biol*, 24, 837-45.
- FISHER, T. S., TAGGART, A. K. & ZAKIAN, V. A. 2004. Cell cycle-dependent regulation of yeast telomerase by Ku. *Nat Struct Mol Biol*, 11, 1198-205.



- FLEISCHER, T. C., WEAVER, C. M., MCAFEE, K. J., JENNINGS, J. L. & LINK, A. J. 2006a. Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. *Genes Dev*, 20, 1294-307.
- FLEISCHER, T. C., WEAVER, C. M., MCAFEE, K. J., JENNINGS, J. L. & LINK, A. J. 2006b. Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. *Genes & development*, 20, 1294-1307.
- FOSTER, S. S., ZUBKO, M. K., GUILLARD, S. & LYDALL, D. 2006. MRX protects telomeric DNA at uncapped telomeres of budding yeast cdc13-1 mutants. *DNA repair*, 5, 840-851.
- GABA, A., JACOBSON, A. & SACHS, M. S. 2005a. Ribosome occupancy of the yeast CPA1 upstream open reading frame termination codon modulates nonsense-mediated mRNA decay. *Mol Cell*, 20, 449-60.
- GABA, A., JACOBSON, A. & SACHS, M. S. 2005b. Ribosome occupancy of the yeast CPA1 upstream open reading frame termination codon modulates nonsense-mediated mRNA decay. *Molecular cell*, 20, 449-460.
- GANDURI, S. & LUE, N. F. 2017. STN1–POLA2 interaction provides a basis for primase-pol  $\alpha$  stimulation by human STN1. *Nucleic Acids Research*, 45, 9455-9466.
- GARDNER, R., PUTNAM, C. W. & WEINERT, T. 1999. RAD53, DUN1 and PDS1 define two parallel G2/M checkpoint pathways in budding yeast. *EMBO J*, 18, 3173-85.
- GARVIK, B., CARSON, M. & HARTWELL, L. 1995. Single-stranded DNA arising at telomeres in cdc13 mutants may constitute a specific signal for the RAD9 checkpoint. *Molecular and Cellular Biology*, 15, 6128-6138.
- GASPARYAN, H. J., XU, L., PETREACA, R. C., REX, A. E., SMALL, V. Y., BHOGAL, N. S., JULIUS, J. A., WARSI, T. H., BACHANT, J., APARICIO, O. M. & NUGENT, C. I. 2009. Yeast telomere capping protein Stn1 overrides DNA replication control through the S phase checkpoint. *Proc Natl Acad Sci U S A*, 106, 2206-11.
- GELINAS, A. D., PASCHINI, M., REYES, F. E., HÉROUX, A., BATEY, R. T., LUNDBLAD, V. & WUTTKE, D. S. 2009. Telomere capping proteins are structurally related to RPA with an additional telomere-specific domain. *Proceedings of the National Academy of Sciences*, 106, 19298-19303.
- GHOSH, S., GANESAN, R., AMRANI, N. & JACOBSON, A. 2010. Translational competence of ribosomes released from a premature termination codon is modulated by NMD factors. *RNA*, 16, 1832-47.
- GIBSON, D. G., YOUNG, L., CHUANG, R. Y., VENTER, J. C., HUTCHISON, C. A., 3RD & SMITH, H. O. 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*, 6, 343-5.
- GILBERT, W. V., ZHOU, K., BUTLER, T. K. & DOUDNA, J. A. 2007. Cap-independent translation is required for starvation-induced differentiation in yeast. *Science*, 317, 1224-7.
- GOCHA, A. R., HARRIS, J. & GRODEN, J. 2013. Alternative mechanisms of telomere lengthening: permissive mutations, DNA repair proteins and tumorigenic progression. *Mutat Res*, 743-744, 142-50.
- GRANDIN, N., DAMON, C. & CHARBONNEAU, M. 2001a. Cdc13 prevents telomere uncapping and Rad50-dependent homologous recombination. *EMBO J*, 20, 6127-39.
- GRANDIN, N., DAMON, C. & CHARBONNEAU, M. 2001b. Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13. *EMBO J*, 20, 1173-83.
- GRANDIN, N., REED, S. I. & CHARBONNEAU, M. 1997. Stn1, a new *Saccharomyces cerevisiae* protein, is implicated in telomere size regulation in association with Cdc13. *Genes Dev*, 11, 512-27.
- GREENE, L. E., PARK, Y. N., MASISON, D. C. & EISENBERG, E. 2009. Application of GFP-labeling to study prions in yeast. *Protein Pept Lett*, 16, 635-41.
- GUNIŠOVÁ, S. & VALÁŠEK, L. S. 2014. Fail-safe mechanism of GCN4 translational control—uORF2 promotes reinitiation by analogous mechanism to uORF1 and thus secures its key role in GCN4 expression. *Nucleic acids research*, gku204.
- HEISS, N. S., KNIGHT, S. W., VULLIAMY, T. J., KLAUCK, S. M., WIEMANN, S., MASON, P. J., POUSTKA, A. & DOKAL, I. 1998. X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. *Nature genetics*, 19, 32-38.

- HERBERT, G. B., SHI, B. & GARTENHAUS, R. B. 2001. Expression and stabilization of the MCT-1 protein by DNA damaging agents. *Oncogene*, 20, 6777-83.
- HINNEBUSCH, A. G. 2005. Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol*, 59, 407-50.
- HINNEBUSCH, A. G., DEVER, T. E. & ASANO, K. 2007. 9 Mechanism of Translation Initiation in the Yeast *Saccharomyces cerevisiae*. *Cold Spring Harbor Monograph Archive*, 48, 225-268.
- HOLSTEIN, E. M., CLARK, K. R. & LYDALL, D. 2014. Interplay between nonsense-mediated mRNA decay and DNA damage response pathways reveals that Stn1 and Ten1 are the key CST telomere-cap components. *Cell Rep*, 7, 1259-69.
- HOLSTEIN, E. M., NGO, G., LAWLESS, C., BANKS, P., GREETHAM, M., WILKINSON, D. & LYDALL, D. 2017. Systematic Analysis of the DNA Damage Response Network in Telomere Defective Budding Yeast. *G3 (Bethesda)*, 7, 2375-2389.
- HSU, H. L., CHOY, C. O., KASIAPPAN, R., SHIH, H. J., SAWYER, J. R., SHU, C. L., CHU, K. L., CHEN, Y. R., HSU, H. F. & GARTENHAUS, R. B. 2007. MCT-1 oncogene downregulates p53 and destabilizes genome structure in the response to DNA double-strand damage. *DNA Repair (Amst)*, 6, 1319-32.
- HSU, H. L., SHI, B. & GARTENHAUS, R. B. 2005. The MCT-1 oncogene product impairs cell cycle checkpoint control and transforms human mammary epithelial cells. *Oncogene*, 24, 4956-64.
- IYER, S., CHADHA, A. D. & MCEACHERN, M. J. 2005. A mutation in the STN1 gene triggers an alternative lengthening of telomere-like runaway recombinational telomere elongation and rapid deletion in yeast. *Mol Cell Biol*, 25, 8064-73.
- JACKSON, R. J., HELLEN, C. U. & PESTOVA, T. V. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol*, 11, 113-27.
- JAFRI, M. A., ANSARI, S. A., ALQAHTANI, M. H. & SHAY, J. W. 2016. Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Genome Med*, 8, 69.
- JIVOTOVSKAYA, A. V., VALÁŠEK, L., HINNEBUSCH, A. G. & NIELSEN, K. H. 2006. Eukaryotic translation initiation factor 3 (eIF3) and eIF2 can promote mRNA binding to 40S subunits independently of eIF4G in yeast. *Molecular and cellular biology*, 26, 1355-1372.
- KEELING, K. M., LANIER, J., DU, M., SALAS-MARCO, J., GAO, L., KAENJAK-ANGELETTI, A. & BEDWELL, D. M. 2004. Leaky termination at premature stop codons antagonizes nonsense-mediated mRNA decay in *S. cerevisiae*. *RNA*, 10, 691-703.
- KERVESTIN, S. & JACOBSON, A. 2012. NMD: a multifaceted response to premature translational termination. *Nature Reviews Molecular Cell Biology*, 13, 700-712.
- KOZAK, M. 1986a. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, 44, 283-292.
- KOZAK, M. 1986b. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, 44, 283-92.
- KUPIEC, M. 2014. Biology of telomeres: lessons from budding yeast. *FEMS microbiology reviews*, 38, 144-171.
- LARRIVÉE, M., LEBEL, C. & WELLINGER, R. J. 2004. The generation of proper constitutive G-tails on yeast telomeres is dependent on the MRX complex. *Genes & development*, 18, 1391-1396.
- LARRIVÉE, M. & WELLINGER, R. J. 2006. Telomerase- and capping-independent yeast survivors with alternate telomere states. *Nat Cell Biol*, 8, 741-7.
- LAWLESS, C., PEARSON, R. D., SELLEY, J. N., SMIRNOVA, J. B., GRANT, C. M., ASHE, M. P., PAVITT, G. D. & HUBBARD, S. J. 2009. Upstream sequence elements direct post-transcriptional regulation of gene expression under stress conditions in yeast. *BMC Genomics*, 10, 7.
- LEE, J. H., PESTOVA, T. V., SHIN, B.-S., CAO, C., CHOI, S. K. & DEVER, T. E. 2002. Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation. *Proceedings of the National Academy of Sciences*, 99, 16689-16694.
- LEI, M., PODELL, E. R. & CECH, T. R. 2004. Structure of human POT1 bound to telomeric single-stranded DNA provides a model for chromosome end-protection. *Nat Struct Mol Biol*, 11, 1223-9.

- LINGNER, J., CECH, T. R., HUGHES, T. R. & LUNDBLAD, V. 1997a. Three Ever Shorter Telomere (EST) genes are dispensable for in vitro yeast telomerase activity. *Proc Natl Acad Sci U S A*, 94, 11190-5.
- LINGNER, J., HUGHES, T. R., SHEVCHENKO, A., MANN, M., LUNDBLAD, V. & CECH, T. R. 1997b. Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science*, 276, 561-7.
- LLOYD, S. E., MEAD, S. & COLLINGE, J. 2013. Genetics of prion diseases. *Curr Opin Genet Dev*, 23, 345-51.
- LOMAKIN, I. B., STOLBOUSHKINA, E. A., VAIDYA, A. T., ZHAO, C., GARBER, M. B., DMITRIEV, S. E. & STEITZ, T. A. 2017. Crystal Structure of the Human Ribosome in Complex with DENR-MCT-1. *Cell Rep*, 20, 521-528.
- LONGTINE, M. S., MCKENZIE, A., 3RD, DEMARINI, D. J., SHAH, N. G., WACH, A., BRACHAT, A., PHILIPPSEN, P. & PRINGLE, J. R. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast*, 14, 953-61.
- LUE, N. F., CHAN, J., WRIGHT, W. E. & HURWITZ, J. 2014. The CDC13-STN1-TEN1 complex stimulates Pol alpha activity by promoting RNA priming and primase-to-polymerase switch. *Nat Commun*, 5, 5762.
- LUKE, B., AZZALIN, C. M., HUG, N., DEPLAZES, A., PETER, M. & LINGNER, J. 2007. *Saccharomyces cerevisiae* Ebs1p is a putative ortholog of human Smg7 and promotes nonsense-mediated mRNA decay. *Nucleic Acids Res*, 35, 7688-97.
- LYDALL, D. 2009. Taming the tiger by the tail: modulation of DNA damage responses by telomeres. *EMBO J*, 28, 2174-87.
- MAAG, D., FEKETE, C. A., GRYCZYNSKI, Z. & LORSCH, J. R. 2005. A conformational change in the eukaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon. *Molecular cell*, 17, 265-275.
- MAJKA, J., NIEDZIELA-MAJKA, A. & BURGERS, P. M. 2006. The checkpoint clamp activates Mec1 kinase during initiation of the DNA damage checkpoint. *Molecular cell*, 24, 891-901.
- MARCAND, S., GILSON, E. & SHORE, D. 1997. A protein-counting mechanism for telomere length regulation in yeast. *Science*, 275, 986-90.
- MARCAND, S., PARDO, B., GRATIAS, A., CAHUN, S. & CALLEBAUT, I. 2008. Multiple pathways inhibit NHEJ at telomeres. *Genes Dev*, 22, 1153-8.
- MARINGELE, L. & LYDALL, D. 2002a. EXO1-dependent single-stranded DNA at telomeres activates subsets of DNA damage and spindle checkpoint pathways in budding yeast yku70Delta mutants. *Genes Dev*, 16, 1919-33.
- MARINGELE, L. & LYDALL, D. 2002b. EXO1-dependent single-stranded DNA at telomeres activates subsets of DNA damage and spindle checkpoint pathways in budding yeast yku70Delta mutants. *Genes & development*, 16, 1919-1933.
- MCEACHERN, M. J. & BLACKBURN, E. H. 1994. A conserved sequence motif within the exceptionally diverse telomeric sequences of budding yeasts. *Proceedings of the National Academy of Sciences*, 91, 3453-3457.
- MITCHELL, P. & TOLLERVEY, D. 2003. An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3'-->5' degradation. *Mol Cell*, 11, 1405-13.
- MONDOUX, M. A. & ZAKIAN, V. A. 2007. Subtelomeric elements influence but do not determine silencing levels at *Saccharomyces cerevisiae* telomeres. *Genetics*, 177, 2541-2546.
- MUNZAROVA, V., PANEK, J., GUNISOVA, S., DANYI, I., SZAMECZ, B. & VALASEK, L. S. 2011. Translation reinitiation relies on the interaction between eIF3a/TIF32 and progressively folded cis-acting mRNA elements preceding short uORFs. *PLoS Genet*, 7, e1002137.
- NAGALAKSHMI, U., WANG, Z., WAERN, K., SHOU, C., RAHA, D., GERSTEIN, M. & SNYDER, M. 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science*, 320, 1344-9.
- NANDI, S., REINERT, L. S., HACHEM, A., MAZAN-MAMCZARZ, K., HAGNER, P., HE, H. & GARTENHAUS, R. B. 2007. Phosphorylation of MCT-1 by p44/42 MAPK is required for its stabilization in response to DNA damage. *Oncogene*, 26, 2283-9.

- NGO, G. H. P., BALAKRISHNAN, L., DUBARRY, M., CAMPBELL, J. L. & LYDALL, D. 2014. The 9-1-1 checkpoint clamp stimulates DNA resection by Dna2-Sgs1 and Exo1. *Nucleic acids research*, gku746.
- NUGENT, C. I., HUGHES, T. R., LUE, N. F. & LUNDBLAD, V. 1996a. Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science*, 274, 249-252.
- NUGENT, C. I., HUGHES, T. R., LUE, N. F. & LUNDBLAD, V. 1996b. Cdc13p: a single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science*, 274, 249-52.
- OLOVNIKOV, A. M. 1973. A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *J Theor Biol*, 41, 181-90.
- PAZ, I., ABRAMOVITZ, L. & CHODER, M. 1999. Starved *Saccharomyces cerevisiae* cells have the capacity to support internal initiation of translation. *Journal of Biological Chemistry*, 274, 21741-21745.
- PELECHANO, V., WEI, W. & STEINMETZ, L. M. 2013. Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature*, 497, 127-31.
- PENNOCK, E., BUCKLEY, K. & LUNDBLAD, V. 2001. Cdc13 delivers separate complexes to the telomere for end protection and replication. *Cell*, 104, 387-96.
- PESTOVA, T. V. & KOLUPAEVA, V. G. 2002. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes & development*, 16, 2906-2922.
- PETERSON, S. E., STELLWAGEN, A. E., DIEDE, S. J., SINGER, M. S., HAIMBERGER, Z. W., JOHNSON, C. O., TZONEVA, M. & GOTTSCHLING, D. E. 2001. The function of a stem-loop in telomerase RNA is linked to the DNA repair protein Ku. *Nat Genet*, 27, 64-7.
- PISAREV, A. V., HELLEN, C. U. & PESTOVA, T. V. 2007. Recycling of eukaryotic posttermination ribosomal complexes. *Cell*, 131, 286-99.
- PISAREV, A. V., SKABKIN, M. A., PISAREVA, V. P., SKABKINA, O. V., RAKOTONDRAFARA, A. M., HENTZE, M. W., HELLEN, C. U. & PESTOVA, T. V. 2010. The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Mol Cell*, 37, 196-210.
- POWLEY, I. R., KONDRASHOV, A., YOUNG, L. A., DOBBYN, H. C., HILL, K., CANNELL, I. G., STONELEY, M., KONG, Y. W., COTES, J. A., SMITH, G. C., WEK, R., HAYES, C., GANT, T. W., SPRIGGS, K. A., BUSHELL, M. & WILLIS, A. E. 2009. Translational reprogramming following UVB irradiation is mediated by DNA-PKcs and allows selective recruitment to the polysomes of mRNAs encoding DNA repair enzymes. *Genes Dev*, 23, 1207-20.
- PRICE, C. M., BOLTZ, K. A., CHAIKEN, M. F., STEWART, J. A., BEILSTEIN, M. A. & SHIPPEN, D. E. 2010. Evolution of CST function in telomere maintenance. *Cell Cycle*, 9, 3157-65.
- PTUSHKINA, M., VON DER HAAR, T., VASILESCU, S., FRANK, R., BIRKENHÄGER, R. & MCCARTHY, J. E. G. 1998. Cooperative modulation by eIF4G of eIF4E-binding to the mRNA 5' cap in yeast involves a site partially shared by p20. *The EMBO journal*, 17, 4798-4808.
- PUGLISI, A., BIANCHI, A., LEMMENS, L., DAMAY, P. & SHORE, D. 2008. Distinct roles for yeast Stn1 in telomere capping and telomerase inhibition. *EMBO J*, 27, 2328-39.
- QI, H. & ZAKIAN, V. A. 2000. The *Saccharomyces* telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase alpha and the telomerase-associated est1 protein. *Genes Dev*, 14, 1777-88.
- RAJAGOPAL, V., PARK, E.-H., HINNEBUSCH, A. G. & LORSCH, J. R. 2012. Specific domains in yeast translation initiation factor eIF4G strongly bias RNA unwinding activity of the eIF4F complex toward duplexes with 5'-overhangs. *Journal of Biological Chemistry*, 287, 20301-20312.
- RAJKOWITSCH, L., VILELA, C., BERTHELOT, K., RAMIREZ, C. V. & MCCARTHY, J. E. 2004. Reinitiation and recycling are distinct processes occurring downstream of translation termination in yeast. *J Mol Biol*, 335, 71-85.
- REINERT, L. S., SHI, B., NANDI, S., MAZAN-MAMCZARZ, K., VITOLO, M., BACHMAN, K. E., HE, H. & GARTENHAUS, R. B. 2006. MCT-1 protein interacts with the cap complex and modulates messenger RNA translational profiles. *Cancer research*, 66, 8994-9001.

- ROY, B., VAUGHN, J. N., KIM, B. H., ZHOU, F., GILCHRIST, M. A. & VON ARNIM, A. G. 2010. The h subunit of eIF3 promotes reinitiation competence during translation of mRNAs harboring upstream open reading frames. *RNA*, 16, 748-61.
- RUGGERO, D. & PANDOLFI, P. P. 2003. Does the ribosome translate cancer? *Nature Reviews Cancer*, 3, 179-192.
- SCHLEICH, S., ACEVEDO, J. M., CLEMM VON HOHENBERG, K. & TELEMAN, A. A. 2017. Identification of transcripts with short stuORFs as targets for DENR-MCTS1-dependent translation in human cells. *Sci Rep*, 7, 3722.
- SCHLEICH, S., STRASSBURGER, K., JANIESCH, P. C., KOLEDACHKINA, T., MILLER, K. K., HANEKE, K., CHENG, Y.-S., KÜCHLER, K., STOECKLIN, G. & DUNCAN, K. E. 2014a. DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth. *Nature*, 512, 208-212.
- SCHLEICH, S., STRASSBURGER, K., JANIESCH, P. C., KOLEDACHKINA, T., MILLER, K. K., HANEKE, K., CHENG, Y. S., KUCHLER, K., STOECKLIN, G., DUNCAN, K. E. & TELEMAN, A. A. 2014b. DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth. *Nature*, 512, 208-12.
- SERIO, T. R. & LINDQUIST, S. L. 1999. [PSI<sup>+</sup>]: an epigenetic modulator of translation termination efficiency. *Annu Rev Cell Dev Biol*, 15, 661-703.
- SHOEMAKER, C. J. & GREEN, R. 2011. Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proc Natl Acad Sci U S A*, 108, E1392-8.
- SHOR, E., GANGLOFF, S., WAGNER, M., WEINSTEIN, J., PRICE, G. & ROTHSTEIN, R. 2002. Mutations in homologous recombination genes rescue top3 slow growth in *Saccharomyces cerevisiae*. *Genetics*, 162, 647-62.
- SIMON, A. J., LEV, A., ZHANG, Y., WEISS, B., RYLOVA, A., EYAL, E., KOL, N., BAREL, O., CESARKAS, K., SOUDACK, M., GREENBERG-KUSHNIR, N., RHODES, M., WIEST, D. L., SCHIBY, G., BARSHACK, I., KATZ, S., PRAS, E., PORAN, H., REZNIK-WOLF, H., RIBAKOVSKY, E., SIMON, C., HAZOU, W., SIDI, Y., LAHAD, A., KATZIR, H., SAGIE, S., AQEILAN, H. A., GLOUSKER, G., AMARIGLIO, N., TZFATI, Y., SELIG, S., REHAVI, G. & SOMECH, R. 2016. Mutations in STN1 cause Coats plus syndrome and are associated with genomic and telomere defects. *J Exp Med*, 213, 1429-40.
- SINGER, M. S. & GOTTSCHLING, D. E. 1994. TLC1: template RNA component of *Saccharomyces cerevisiae* telomerase. *Science*, 266, 404-9.
- SKABKIN, M. A., SKABKINA, O. V., DHOTE, V., KOMAR, A. A., HELLEN, C. U. & PESTOVA, T. V. 2010a. Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Genes Dev*, 24, 1787-801.
- SKABKIN, M. A., SKABKINA, O. V., DHOTE, V., KOMAR, A. A., HELLEN, C. U. & PESTOVA, T. V. 2010b. Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Genes & development*, 24, 1787-1801.
- SOMERS, J., WILSON, L. A., KILDAY, J. P., HORVILLEUR, E., CANNELL, I. G., POYRY, T. A., COBBOLD, L. C., KONDRASHOV, A., KNIGHT, J. R., PUGET, S., GRILL, J., GRUNDY, R. G., BUSHELL, M. & WILLIS, A. E. 2015. A common polymorphism in the 5' UTR of ERCC5 creates an upstream ORF that confers resistance to platinum-based chemotherapy. *Genes Dev*, 29, 1891-6.
- STELLWAGEN, A. E., HAIMBERGER, Z. W., VEATCH, J. R. & GOTTSCHLING, D. E. 2003. Ku interacts with telomerase RNA to promote telomere addition at native and broken chromosome ends. *Genes Dev*, 17, 2384-95.
- SUN, J., YU, E. Y., YANG, Y., CONFER, L. A., SUN, S. H., WAN, K., LUE, N. F. & LEI, M. 2009. Stn1-Ten1 is an Rpa2-Rpa3-like complex at telomeres. *Genes Dev*, 23, 2900-14.
- SWISHER, K. D. & PARKER, R. 2011. Interactions between Upf1 and the decapping factors Edc3 and Pat1 in *Saccharomyces cerevisiae*. *PLoS One*, 6, e26547.
- TANNER, N. K. & LINDER, P. 2001. DEXD/H box RNA helicases: from generic motors to specific dissociation functions. *Molecular cell*, 8, 251-262.
- TARUN JR, S. Z. & SACHS, A. B. 1996. Association of the yeast poly (A) tail binding protein with translation initiation factor eIF-4G. *The EMBO journal*, 15, 7168.

- VIALARD, J. E., GILBERT, C. S., GREEN, C. M. & LOWNDES, N. F. 1998a. The budding yeast Rad9 checkpoint protein is subjected to Mec1/Tel1-dependent hyperphosphorylation and interacts with Rad53 after DNA damage. *EMBO J*, 17, 5679-88.
- VIALARD, J. E., GILBERT, C. S., GREEN, C. M. & LOWNDES, N. F. 1998b. The budding yeast Rad9 checkpoint protein is subjected to Mec1/Tel1-dependent hyperphosphorylation and interacts with Rad53 after DNA damage. *The EMBO Journal*, 17, 5679-5688.
- WALKER, S. E., ZHOU, F., MITCHELL, S. F., LARSON, V. S., VALASEK, L., HINNEBUSCH, A. G. & LORSCH, J. R. 2013. Yeast eIF4B binds to the head of the 40S ribosomal subunit and promotes mRNA recruitment through its N-terminal and internal repeat domains. *RNA*, 19, 191-207.
- WANG, Z., FANG, P. & SACHS, M. S. 1998. The evolutionarily conserved eukaryotic arginine attenuator peptide regulates the movement of ribosomes that have translated it. *Molecular and cellular biology*, 18, 7528-7536.
- WATSON, J. D. 1972. Origin of concatemeric T7 DNA. *Nat New Biol*, 239, 197-201.
- WEBB, C. J., WU, Y. & ZAKIAN, V. A. 2013. DNA repair at telomeres: Keeping the ends intact. *Cold Spring Harbor perspectives in biology*, 5.
- WELLINGER, R. J. 2010. When the caps fall off: responses to telomere uncapping in yeast. *FEBS letters*, 584, 3734-3740.
- WELLINGER, R. J., WOLF, A. J. & ZAKIAN, V. A. 1993. Saccharomyces telomeres acquire single-strand TG 1-3 tails late in S phase. *Cell*, 72, 51-60.
- WETHMAR, K. 2014. The regulatory potential of upstream open reading frames in eukaryotic gene expression. *Wiley Interdiscip Rev RNA*, 5, 765-78.
- YAMAN, I., FERNANDEZ, J., LIU, H., CAPRARA, M., KOMAR, A. A., KOROMILAS, A. E., ZHOU, L., SNIDER, M. D., SCHEUNER, D., KAUFMAN, R. J. & HATZOGLOU, M. 2003. The zipper model of translational control: a small upstream ORF is the switch that controls structural remodeling of an mRNA leader. *Cell*, 113, 519-31.
- YOUNG, S. K. & WEK, R. C. 2016. Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. *J Biol Chem*, 291, 16927-35.
- ZABORSKE, J. M., ZEITLER, B. & CULBERTSON, M. R. 2013. Multiple Transcripts from a 3'-UTR Reporter Vary in Sensitivity to Nonsense-Mediated mRNA Decay in Saccharomyces cerevisiae. *PLoS one*, 8, e80981.
- ZHOU, Q. & CHAI, W. 2016. Suppression of STN1 enhances the cytotoxicity of chemotherapeutic agents in cancer cells by elevating DNA damage. *Oncol Lett*, 12, 800-808.
- ZINOVIEV, A., HELLEN, C. U. & PESTOVA, T. V. 2015. Multiple mechanisms of reinitiation on bicistronic calicivirus mRNAs. *Mol Cell*, 57, 1059-73.
- ZUBKO, M. K., GUILLARD, S. & LYDALL, D. 2004. Exo1 and Rad24 differentially regulate generation of ssDNA at telomeres of Saccharomyces cerevisiae cdc13-1 mutants. *Genetics*, 168, 103-15.

